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: April 24, 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]

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

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. $\lambda_{\text{SO}} > 0$ 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

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

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

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

The corresponding eigenenergies are

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

with $\Delta = \sqrt{\lambda_{\text{SO}}^2 + 4(\Upsilon_x^2 + \Upsilon_y^2)} \approx 2\pi \times 46$ GHz. Since $\Upsilon_{x,y} \ll \lambda_{\text{SO}}$ (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

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

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 $\ddot{A}_{n,k}(t) + \omega_{n,k}^2 A_{n,k}(t) = 0$, and the mode frequencies $\omega_{n,k}$ and the transverse mode profile $\vec{u}_{n,k}^{\perp}(x, y)$ are in general obtained from a numerical solution of the Eq. (6). The $\vec{u}_{n,k}^{\perp}(x, y)$ are orthogonal and normalized to

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

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:

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

where $M = \rho A L$ 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

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

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

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

in terms of annihilation and creation operators. We obtain

$$H_{\text{ph}} = \sum_{k,n} \hbar \omega_{n,k} a_{n,k}^{\dagger} a_{n,k}, \tag{15}$$

and the quantized displacement field

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

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

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

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

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

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

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

$$
\equiv d\sqrt{\frac{\hbar k^2}{2\rho A \omega_{n,k}}}\xi_{n,k}(y,z). \tag{23}
$$

Here, $u_{n,k}^{\perp,i}$ represents the i-th component of the displacement pattern $\vec{u}_{n,k}^{\perp}(y,z)$. 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,

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

where

$$H_{\text{drive}} = \frac{\Omega(t)}{2} e^{i[\omega_d t + \theta(t)]} (|3\rangle\langle 2| + |4\rangle\langle 1|) + \text{H.c.} \tag{26}$$

By moving into the interaction picture with respect to

$$H_0 = \sum_{n,k} \omega_0 a_{n,k}^\dagger a_{n,k} + \omega_B |2\rangle\langle 2| + \omega_0 |3\rangle\langle 3| + (\omega_0 + \omega_B)|4\rangle\langle 4|, \tag{27}$$

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

$$\tilde{H} = \sum_{n,k} (\omega_{n,k} - \omega_0) a_{n,k}^\dagger a_{n,k} - \delta(|3\rangle\langle 3| + |4\rangle\langle 4|)$$
$$+ \left[\frac{\Omega(t) e^{i\theta(t)}}{2} (|3\rangle\langle 2| + e^{2i\omega_B t}|4\rangle\langle 1|) + \frac{1}{\sqrt{L}} \sum_{n,k} g_{n,k} e^{ikx} a_{n,k} (|3\rangle\langle 1| + |4\rangle\langle 2|) + \text{H.c.} \right]. \tag{28}$$

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

$$|\psi(t)\rangle = \alpha|1,0\rangle + \beta \left[c(t)|2\rangle\langle 1| + b(t)|3\rangle\langle 1| + \sum_{n,k} c_{n,k}(t) a_{n,k}^\dagger \right]|1,0\rangle, \tag{29}$$

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 $|1\rangle \rightarrow |4\rangle$ produced by the drive [i.e., the term $\sim e^{2i\omega_B t}$ in Eq. (28)]. 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,

$$\dot{c}_{n,k}(t) = -i(\omega_{n,k} - \omega_0)c_{n,k}(t) - i\frac{1}{\sqrt{L}}g^*_{n,k}e^{-ikx}b(t),$$

$$\dot{c}(t) = -i\frac{\Omega(t)e^{-i\theta(t)}}{2}b(t),$$ (30)

$$\dot{b}(t) = i\delta b(t) - i\frac{\Omega(t)e^{i\theta(t)}}{2}c(t) - i\frac{1}{\sqrt{L}}\sum_{n,k}g_{n,k}e^{ikx}c_{n,k}(t).$$

The solution for the propagating phonons reads

$$c_{n,k}(t) = e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0) - \frac{i}{\sqrt{L}}g^*_{n,k}e^{-ikx}\int_{t_0}^t d\tau e^{-i(\omega_{n,k}-\omega_0)(t-\tau)}b(\tau),$$ (31)

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

$$\dot{b}(t) = i\delta b(t) - i\frac{\Omega(t)e^{i\theta(t)}}{2}c(t) - i\frac{1}{\sqrt{L}}\sum_{n,k}g_{n,k}e^{ikx}e^{-i(\omega_{n,k}-\omega_0)t}c_{n,k}(t_0) - \frac{1}{L}\sum_{n,k}|g_{n,k}|^2\int_{t_0}^t d\tau e^{-i(\omega_{n,k}-\omega_0)(t-\tau)}b(\tau).$$ (32)

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

$$\dot{b}(t) = \left[i\delta - \frac{\Gamma(\omega_0)}{2}\right]b(t) - i\frac{\Omega(t)e^{i\theta(t)}}{2}c(t) - \sum_n\sqrt{\frac{\Gamma_n(\omega_0)}{2}}\left[\Phi^{\text{in,L}}_{n,\omega_0}(t) + \Phi^{\text{in,R}}_{n,\omega_0}(t)\right],$$ (33)

with the phonon-induced decay rate

$$\Gamma(\omega) = \frac{2\pi}{L}\sum_{n,k}|g_{n,k}|^2\delta(\omega - \omega_{n,k}) = 2\frac{|g_n|^2}{v_n},$$ (34)

and the input field

$$\Phi^{\text{in,L/R}}_{n,\omega}(t) = i\sum_{k>0}\sqrt{\frac{v_n}{L}}e^{\mp ikx}e^{-i(\omega_{n,k}-\omega)(t-t_0)}c_{n,k}(t_0).$$ (35)

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

$$\dot{c}(t) = -\left[i\omega_s(t) + \frac{\gamma(t)}{2}\right]c(t) - \sum_n\sqrt{\frac{\gamma_n(t)}{2}}e^{-i\bar{\theta}(t)}\left[\Phi^{\text{in,L}}_n(t) + \Phi^{\text{in,R}}_n(t)\right],$$ (36)

with

$$\omega_s(t) = \frac{\Omega^2(t)}{4}\frac{\delta}{\delta^2 + \Gamma^2(\omega_0)/4}, \quad \gamma(t) = \sum_n\gamma_n(t) = \frac{\Omega^2(t)/4}{\delta^2 + \Gamma^2(\omega_0)/4}\sum_n\Gamma_n(\omega_0), \quad \bar{\theta}(t) = \theta(t) + \arctan\left[\frac{\Gamma(\omega_0)}{2\delta}\right]$$ (37)

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.
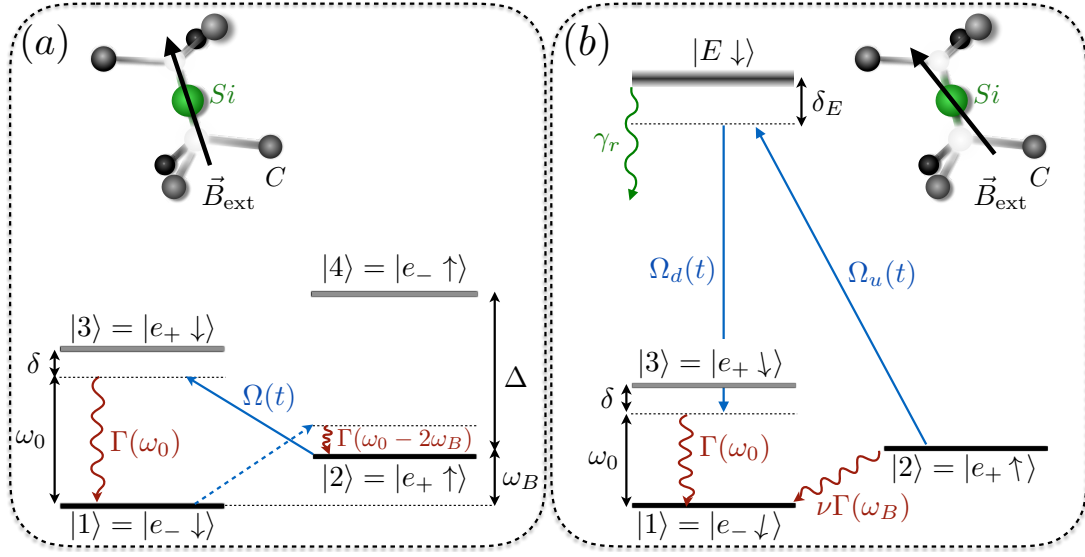
FIG. 1: Comparison of the two driving schemes to implement the spin-phonon interface. (a) A microwave magnetic field drives the transition $|2\rangle \to |3\rangle$, and state $|3\rangle$ subsequently decays to state $|1\rangle$ by emitting a propagating phonon at frequency ω_0. The process that drives 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 $|2\rangle \to |1\rangle$.

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

As described by Eq.(28), 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

$$|\psi(t)\rangle = \alpha|2,0\rangle + \beta\big[c(t)|1\rangle\langle2| + b(t)|4\rangle\langle2| + \sum_{n,k} c_{n,k}(t)a_{n,k}^\dagger\big]|2,0\rangle, \tag{38}$$

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

$$\dot{c}(t) = -\left[i\tilde{\omega}_s(t) + \frac{\tilde{\gamma}(t)}{2}\right]c(t) - \sum_n \sqrt{\frac{\tilde{\gamma}_n(t)}{2}}e^{-i\theta(t)}[\Phi_{n,\omega_0-2\omega_B}^{\text{in,L}}(t) + \Phi_{n,\omega_0-2\omega_B}^{\text{in,R}}(t)], \tag{39}$$

with the input fields defined in Eq. (35), the AC-Stark-shift and the effective transfer rate

$$\tilde{\omega}_s(t) = \frac{\Omega^2(t)}{4}\frac{\delta - 2\omega_B}{(\delta - 2\omega_B)^2 + \Gamma^2(\omega_0 - 2\omega_B)/4}, \qquad \tilde{\gamma}(t) = \frac{\Omega^2(t)/4}{(\delta - 2\omega_B)^2 + \Gamma^2(\omega_0 - 2\omega_B)/4}\sum_n \Gamma_n(\omega_0 - 2\omega_B). \tag{40}$$

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

$$\frac{\tilde{\Gamma}(t)}{\Gamma(t)} \sim \frac{\delta^2}{(\delta - 2\omega_B)^2}\frac{\Gamma_{\text{ph}}(\omega_0 - 2\omega_B)}{\Gamma_{\text{ph}}(\omega_0)} \tag{41}$$

is strongly suppressed, as long as $\Gamma_{\text{ph}}(\omega_0 - 2\omega_B) \simeq \Gamma_{\text{ph}}(\omega_0)$. 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

$$H = H_{\text{SiV}} + H_{\text{ph}} + H_{\text{drive}} + H_{\text{strain}} + H_{\text{rad}}, \tag{42}$$

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].

<div align="center">Effects of a weak perpendicular magnetic field</div>

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

$$
\begin{aligned}
H_{\mathrm{SiV}} &= -\lambda_{\mathrm{SO}} L_z S_z + \omega_B S_z + \omega_x B_x S_x, \\
&= -\frac{\Delta + \omega_B}{2}|e_-\downarrow\rangle\langle e_-\downarrow| - \frac{\Delta - \omega_B}{2}|e_+\uparrow\rangle\langle e_+\uparrow| \\
&\quad + \frac{\Delta - \omega_B}{2}|e_+\downarrow\rangle\langle e_+\downarrow| + \frac{\Delta + \omega_B}{2}|e_-\uparrow\rangle\langle e_-\uparrow| + \omega_E|E\downarrow\rangle\langle E\downarrow|,
\end{aligned}
\tag{43}
$$

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. (43) 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:

$$
H_{\mathrm{SiV}} = \sum_{i=1,4}\omega_i|i\rangle\langle i| + \omega_E|E\downarrow\rangle\langle E\downarrow|, \quad \Rightarrow \quad
\begin{cases}
|1\rangle \approx |e_-\downarrow\rangle - \eta_+|e_-\uparrow\rangle, & \omega_1 \approx -\frac{\Delta + \omega_B}{2} - \frac{\eta_+\omega_x}{2} \\
|2\rangle \approx |e_+\uparrow\rangle - \eta_-|e_+\downarrow\rangle, & \omega_2 \approx -\frac{\Delta - \omega_B}{2} - \frac{\eta_-\omega_x}{2} \\
|3\rangle \approx |e_+\downarrow\rangle + \eta_-|e_+\uparrow\rangle, & \omega_3 \approx \frac{\Delta - \omega_B}{2} + \frac{\eta_-\omega_x}{2} \\
|4\rangle \approx |e_-\uparrow\rangle + \eta_+|e_-\downarrow\rangle, & \omega_4 \approx \frac{\Delta + \omega_B}{2} + \frac{\eta_+\omega_x}{2}
\end{cases}
\tag{44}
$$

with

$$
\eta_\pm \equiv \frac{1}{2}\frac{\omega_x}{\Delta \pm \omega_B}, \qquad \eta = \eta_- + \eta_+.
\tag{45}
$$

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

$$
H_{\mathrm{strain}} = \sum_{n,k} g_{n,k}a_{n,k}\left[J_+ + \eta(|4\rangle\langle 3| - |2\rangle\langle 1|)\right] + \mathrm{H.c.}
\tag{46}
$$

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

$$
\begin{aligned}
H_{\mathrm{drive}} &= \left(\frac{\Omega_d(t)e^{i\theta_d(t)}}{2}e^{-i\omega_d t} + \frac{\Omega_u(t)e^{i\theta_u(t)}}{2}e^{-i\omega_u t}\right)|E\downarrow\rangle\langle e_+\downarrow| + \mathrm{H.c.} \\
&= \frac{\Omega_d(t)e^{i\theta_d(t)}}{2}e^{-i\omega_d t}|E\downarrow\rangle\langle 3| - \frac{\Omega_u(t)e^{i\theta_u(t)}}{2}\eta_- e^{-i\omega_u t}|E\downarrow\rangle\langle 2| + \mathrm{H.c.}
\end{aligned}
\tag{47}
$$

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

<div align="center">Effective 3-level system</div>

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

$$
H_0 = \sum_{n,k}\omega_0 a_{n,k}^\dagger a_{n,k} + \sum_i \omega_i|i\rangle\langle i| + \delta(|3\rangle\langle 3| + |4\rangle\langle 4|) + \delta_E|E\downarrow\rangle\langle E\downarrow|,
\tag{48}
$$

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

$$|\psi(t)\rangle = \alpha|1,0\rangle + \beta\big[c(t)|2\rangle\langle 1| + b(t)|3\rangle\langle 1| + E(t)|E\downarrow\rangle\langle 1| + \sum_{n,k} c_{n,k}(t)a_{n,k}^\dagger\big]|1,0\rangle, \tag{49}$$

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.

$$\dot{E}(t) = \left(i\delta_E - \frac{\Gamma_{\rm rad}}{2}\right)E(t) - \frac{i}{2}\big[\Omega_d(t)e^{i\theta_d(t)}b(t) - \eta_-\Omega_u(t)e^{i\theta_u(t)}c(t)\big]. \tag{50}$$

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

$$\dot{b}(t) = -\left\{i\big[\omega_d(t) - \delta\big] + \frac{\gamma_{\rm rad}^d(t)}{2} + \frac{\Gamma(\omega_0)}{2}\right\}b(t) + i\frac{\Omega_{\rm eff}(t)}{2}e^{i[\theta_{\rm eff}(t) - \phi_{\rm NH}]}c(t) + \sqrt{\frac{\Gamma_n(\omega_0)}{2}}\big[\Phi_{n,\omega_0}^{\rm in,L}(t) + \Phi_{n,\omega_0}^{\rm in,R}(t)\big],$$

$$\dot{c}(t) = -\left\{i\omega_u(t) + \frac{\gamma_{\rm rad}^u(t)}{2} + \frac{\eta^2\Gamma(\omega_B)}{2}\right\}c(t) + i\frac{\Omega_{\rm eff}(t)}{2}e^{-i[\theta_{\rm eff}(t) + \phi_{\rm NH}]}b(t) + \sqrt{\frac{\eta^2\Gamma_n(\omega_B)}{2}}\big[\Phi_{n,\omega_B}^{\rm in,L}(t) + \Phi_{n,\omega_B}^{\rm in,R}(t)\big], \tag{51}$$

with the phonon-induced decay rate and input fields defined in Eqs. (34) and (35) 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. (51), 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 $|3\rangle$,

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

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

$$\gamma_{\rm rad}^u(t) = \frac{\eta_-^2}{4}\frac{\Omega_u^2(t)}{\delta_E^2 + \Gamma_{\rm rad}^2/4}\Gamma_{\rm rad}, \qquad \gamma_{\rm rad}^d(t) = \frac{1}{4}\frac{\Omega_d^2(t)}{\delta_E^2 + \Gamma_{\rm rad}^2/4}\Gamma_{\rm rad}, \tag{53}$$

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 $\Phi_{\omega_B}^{\rm in}(t)$. Finally, the effective Rabi frequency driving the transition $|2\rangle \to |3\rangle$ is given by

$$\Omega_{\rm eff}(t) = \frac{\eta_-}{2}\frac{\Omega_d(t)\Omega_u(t)}{\sqrt{\delta_E^2 + \Gamma_{\rm rad}^2/4}}, \qquad \theta_{\rm eff}(t) = \theta_u(t) - \theta_d(t). \tag{54}$$

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. $\Gamma(\omega_0) \gg \gamma_{\rm rad}^d(t)$, while the other loss mechanisms as to be overcome by the final spin-state transfer rate, i.e. $\gamma(t) \sim \frac{\Omega_{\rm eff}^2(t)}{\delta^2}\Gamma(\omega_0) \gg \gamma_{\rm rad}^u(t), \eta^2\Gamma(\omega_B)$. 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

$$\gamma_{\rm rad}^d \sim 150\,{\rm kHz} \ll \Gamma(\omega_0) \sim 2\,{\rm MHz}, \qquad \gamma_{\rm rad}^u \sim 7\,{\rm kHz}, \eta^2\Gamma(\omega_B) \sim 40\,{\rm kHz} \ll \gamma \sim 250\,{\rm kHz}. \tag{55}$$

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.

Residual driving of neighbouring defects

Both driving schemes discussed above are meant to control individual SiV centers while avoiding any residual driving of defects not involved in the state transfer. For optical control, a natural limit for the separation between SiV centers arises from the diffraction-limited laser spot, i.e. the minimal distance between two consecutive centres should be much larger than the optical wavelength. The optical transition wavelength in the case of the SiV is around 738 nm, meaning that separations $d \sim 10\mu m$ (as in the main text) is sufficient to allow individual control of single defects. This is consistent with, e.g., trapped ion experiments, where the optical control of individual ions separated by $5 - 10$ μm is routinely used in the lab. Note, however, that the diffraction limit can be overcome by various nonlinear techniques; the control of solid-spin states with about 50 nm resolution has been demonstrated in Ref. [9], and similar schemes could potentially be used in the current context.

In the case of microwave magnetic fields, the limitation depends on the specific implementation. As an example, in Ref. [8] the magnetic field is produced by an electric current passing through a thin wire placed at about 200 nm from a spin qubit. In this case the magnetic field approximately falls as $B \sim 1/r$, meaning that the field seen by a neighbouring SiV centre at $d \sim 10\mu m$ far apart would be reduced by a factor 50. Therefore, the resulting effective decay rate, which scales as B^2 (see Eq. 7 of the main text), is a factor of 2500 smaller.

In addition, it could be possible to implement a spatially-dependent static magnetic field so that the Zeeman splitting for two consecutive SiV centres are different [10]. In that case, any unwanted scattering due to residual driving of a neighbouring defect will be further suppressed by the additional mismatch of the Zeeman energies.

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. (27) reads

$$H = \sum_{n,k}(\omega_{n,k} - \omega_0)a_{n_k}^\dagger a_{n_k} + \sum_j H_{\text{SiV}}^{(j)} + \frac{1}{\sqrt{L}}\sum_{j,n,k}\left[g_{n,k}^j e^{ikx_j}a_{n,k}J_+^j + \text{H.c.}\right], \tag{56}$$

with

$$H_{\text{SiV}}^{(j)} = -\delta_j(|3\rangle_j\langle 3| + |4\rangle_j\langle 4|) + \left[\frac{\Omega_j(t)e^{i\theta_j(t)}}{2}|3\rangle_j\langle 2| + \text{H.c.}\right]. \tag{57}$$

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

$$C^\dagger(t) = \sum_{j=e,r}\left[c_j(t)|2\rangle_j\langle 1| + b_j(t)|3\rangle_j\langle 1| + \sum_{n,k}c_{n,k}(t)a_{n,k}^\dagger\right], \tag{58}$$

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. (58), 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

$$\dot{c}_{n,k}(t) = -i(\omega_{n,k} - \omega_0)c_{n,k}(t) - i\frac{1}{\sqrt{L}}(g_{n,k}^e)^* e^{-ikx_e}b_e(t) - i\frac{1}{\sqrt{L}}(g_{n,k}^r)^* e^{-ikx_r}b_r(t),$$

$$\dot{c}_j(t) = -i\frac{\Omega_j(t)e^{-i\theta_j(t)}}{2}b_j(t), \tag{59}$$

$$\dot{b}_j(t) = i\delta_j b_j(t) - i\frac{\Omega_j(t)e^{i\theta_j(t)}}{2}c_j(t) - i\frac{1}{\sqrt{L}}\sum_{n,k}g_{n,k}^j e^{ikx_j}c_{n,k}(t).$$

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

$$\dot{b}_r(t) = \left[i\delta_r - \frac{\Gamma_r(\omega_0)}{2}\right]b_r(t) - i\frac{\Omega_r(t)e^{i\theta_r(t)}}{2}c_r(t) - i\sum_n\sum_k\sqrt{\frac{\Gamma_{r,n}(\omega_0)}{2}}\sqrt{\frac{v_n}{L}}e^{ikx_r}e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0)$$
$$- \sum_n\sqrt{\frac{\Gamma_{r,n}(\omega_0)}{2}\frac{\Gamma_{e,n}(\omega_0)}{2}}e^{ik_n(x_r-x_e)}b_e(t-\tau_{er}^n), \tag{60}$$

with $\tau_{er}^n = (x_r - x_e)/v_n$. 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

$$\dot{c}_r(t) = -\left[i\omega_{s,r} + \frac{\gamma_r(t)}{2}\right]c_r(t) - \sum_n\sqrt{\frac{\gamma_{r,n}(t)}{2}}e^{-i\bar{\theta}_r(t)}[\Phi_{r,n}^{\text{in,L}}(t) + \Phi_{r,n}^{\text{in,R}}(t) + \Phi_{r,n}^{\text{scatt}}(t)]. \tag{61}$$

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. (37) 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)

$$\Phi_{r,n}^{\text{in,L}}(t) = i\sum_{k>0}\sqrt{\frac{v_n}{L}}e^{-ikx_r}e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0),$$
$$\Phi_{r,n}^{\text{in,R}}(t) = i\sum_{k>0}\sqrt{\frac{v_n}{L}}e^{ikx_r}e^{-i(\omega_{n,k}-\omega_0)(t-t_0)}c_{n,k}(t_0) + \sqrt{\frac{\gamma_{e,n}(t-\tau_{er}^n)}{2}}e^{i\theta_e(t-\tau_{er}^n)}e^{ik_n(x_r-x_e)}, \tag{62}$$

while $\Phi_{r,n}^{\text{scatt}}(t)$ 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

$$\Phi_{r,n}^{\text{in,R}}(t) = \Phi_{e,n}^{\text{out,R}}(t-\tau_{er}^n)e^{i\phi_{er}^n} \qquad \therefore \qquad \Phi_{e,n}^{\text{out,R}}(t) = \Phi_{e,n}^{\text{in,R}}(t) + \sqrt{\frac{\gamma_{e,n}(t)}{2}}e^{i\theta_e(t)}, \tag{63}$$

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 $\Phi_{r,n}^{\text{in,L}}(t)$ and $\Phi_{e,n}^{\text{in,R}}(t)$, as described in Eq. (62). 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

$$\Phi_{e,n}^{\text{in,R}}(t) = -\sqrt{R_n}\Phi_{e,n}^{\text{out,L}}(t-\tau_e^n)e^{i\phi_e^n}, \qquad \Phi_{r,n}^{\text{in,L}}(t) = -\sqrt{R_n}\Phi_{r,n}^{\text{out,R}}(t-\tau_r^n)e^{i\phi_r^n}, \tag{64}$$

where the delay times are $\tau_e^n = 2x_e/v_n$, $\tau_r^n = 2(L-x_r)/v_n$ and the phases $\phi_e^n = 2k_nx_e$, $\phi_r^n = 2k_n(L-x_r)$. 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.

$$R_n = e^{-\kappa_n L/v_n}, \qquad \Rightarrow \qquad Q = \frac{\omega_0}{\kappa_n} = -\frac{\omega_0}{\log(R_n)}\frac{L}{v_n}, \quad \kappa_n = -\log(R_n)\frac{v_n}{L} = -\log(R_n)\frac{\Delta\omega_n}{\pi}. \tag{65}$$

For example, $L = 100\,\mu\text{m}$, $v_t = 0.7 \times 10^4$ 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 $Q \approx 4.95 \times 10^4$.

Scattering from undriven centers

As mentioned above, the last term in Eq. (61),

$$\Phi_{r,n}^{\text{scatt}}(t) = -i \sum_{n'} \frac{1}{\delta_e + i\Gamma_e(\omega_0)/2} \sqrt{\frac{\Gamma_{e,n}(\omega_0)}{2} \frac{\Gamma_{e,n'}(\omega_0)}{2}} \Phi_{r,n}^{\text{out,L}}(t - \tau_{er}^n - \tau_{er}^{n'}) e^{i(\phi_{er}^n - \phi_{er}^{n'})}, \tag{66}$$

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 $\delta_j \gg \Gamma_j(\omega_0)$.

Numerical simulation

From the equation of motion [cf. Eq. (61)] (plus the equivalent for the emitting SiV center) and the input-output relations [cf. Eqs. (63) and (64)], it is straightforward to numerically simulate the time evolution of the full Network for a single excitation. To do so, one has to solve the first-order linear differential equation [Eq. (61)] at every time steps (δt later), starting at $t = 0$ with no input fields for both defects. The results presented in this work are obtained using the standard fourth-order Runge-Kutta method.

The drive $\gamma_e(t)$ on the emitting center is fixed while the drive on the receiving center is tuned by minimizing the output field toward the emitting center for a given mode. We chose to minimize the transverse mode (arbitrary choice), so that we tune the drive on the receiving SiV to minimize $\Phi_{r,t}^{\text{out,L}}(t)$ while respecting the constraint on the maximal achievable drive $\gamma_r(t) \le \gamma_{\max}$ at every time step. More precisely, minimizing the output field at time t allows us to determine the drive at time $t + \delta t$. From the solution of both equations, we can compute the outgoing fields $\Phi_{j,n}^{\text{out,R/L}}(t)$ and use them as input fields at later time as prescribed by Eqs. (63) and (64). In our simulations, we neglected scattering event from undriven centres, i.e. $\Phi_{r,n}^{\text{scatt}}(t) = 0$.

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)].

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

$$\vec{u}(\vec{r}) = \sum_{k,n} \sqrt{\frac{\hbar}{M\omega_{n,k}}} \vec{u}_{n,k}^{\perp}(y,z) \big(a_{n,k} + a_{n,k}^{\dagger}\big) \cos(kx), \tag{67}$$
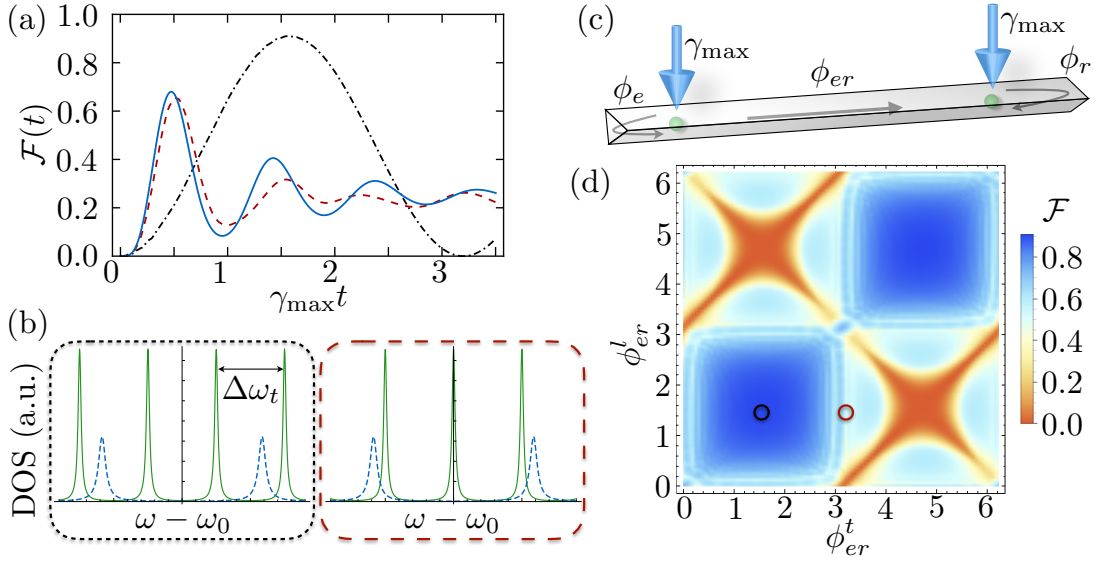
FIG. 2: State transfer for constant driving of the SiV centers. (a) Fidelity as a function of time for two different scenarios. The black dot-dashed line corresponds to both phonon branches being off-resonant, while the dashed red line corresponds to only the transverse being resonant. In the latter 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_{er}^t) and longitudinal (ϕ_{er}^l) 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 ($\phi_e^n = \phi_r^n = \pi$), a frequency splitting $\Delta\omega_t/\gamma_{max} = 140$ and a reflectivity $R = 0.92$. (see main text)

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

$$H_{\text{strain}} = \sqrt{\frac{2}{L}} \sum_{n,k} g_{n,k} \sin(kx) J_+ a_{n,k} + \text{H.c.} \approx \sqrt{\frac{2}{L}} g_{t,k_t} \sin(k_t x) J_+ a + \text{H.c.}, \tag{68}$$

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. (28)], the effective Hamiltonian describing the center as an effective 2-level system coupled to a single phonon mode reads

$$H = \delta a^\dagger a - i g \sigma_+ a + \text{H.c.}, \qquad \text{with} \qquad g = \sqrt{\frac{2}{L}} \frac{g_{t,k_t} \Omega}{2\delta} \sin(k_t x), \tag{69}$$

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

$$H_{\text{s.m.}} = \delta a^\dagger a + g_e \sigma_{e,+} a + g_r \sigma_{r,+} a + \text{H.c.}, \qquad \therefore \qquad g_j = \sqrt{\frac{2}{L}} \frac{g_{t,k_t}^j \Omega_j}{2\delta_j} \sin(k_t x_j). \tag{70}$$

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. (37)]

$$\gamma_{j,t} = \frac{\Omega_j^2/4}{\delta_j^2 + \Gamma_j^2(\omega_0)/4} \Gamma_{j,t}(\omega_0) \approx \frac{\Omega_j^2}{2\delta_j^2} \frac{|g_{t,k_t}^j|^2}{v_t}. \tag{71}$$

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

$$|g_j| = \sqrt{\frac{\gamma_{\max}\Delta\omega_t}{2\pi}}\sin(k_t x_j) = \sqrt{\frac{v_t\gamma_{\max}}{2L}}\sin(k_t x_j), \tag{72}$$

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. (70). Focusing on the low-excitation wavefunction

$$|\psi(t)\rangle = \alpha\mathbb{1}|0\rangle + \sum_{j=e,r}\left[c_j(t)\sigma_{j,+} + c_p(t)a^\dagger\right]|0\rangle, \tag{73}$$

the time evolution is given by

$$\partial_t c_j(t) = -igc_p(t), \qquad \partial_t c_p(t) = -(i\delta - \kappa/2)c_p(t) - ig[c_1(t) + c_2(t)]. \tag{74}$$

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. (65)]. Including the initial conditions $c_e(0) = 1$ and $c_r(0) = c_p(0) = 0$, the solutions read

$$c_e(t) = 1 + c\frac{g}{\tilde{\omega}_-}(e^{-i\tilde{\omega}_-t} - 1) - c\frac{g}{\tilde{\omega}_+}(e^{-i\tilde{\omega}_+t} - 1),$$

$$c_r(t) = -1 + c\frac{g}{\tilde{\omega}_-}(e^{-i\tilde{\omega}_-t} + 1) - c\frac{g}{\tilde{\omega}_+}(e^{-i\tilde{\omega}_+t} + 1), \tag{75}$$

$$c_p(t) = ce^{-i\tilde{\omega}_-t} - ce^{-i\tilde{\omega}_+t},$$

with

$$\tilde{\omega}_\pm = \frac{\delta}{2} - i\frac{\kappa}{4} \pm \sqrt{2g^2 + \frac{1}{4}\left(\delta - i\frac{\kappa}{2}\right)^2}, \qquad \text{and} \qquad c = \frac{1}{2g}\left[\frac{1}{\tilde{\omega}_-} - \frac{1}{\tilde{\omega}_+}\right]^{-1}. \tag{76}$$

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

$$|c_r(t)|^2 \approx \frac{1}{4}[1 - \cos(\sqrt{2}gt)e^{-\kappa t/4}]^2. \tag{77}$$

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 $\Delta\omega_t/\gamma_{\max} = 140$ 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_{er}^n, becomes relevant. In Fig. 2 (d), we plot the state-transfer fidelity as a function of ϕ_{er}^t and ϕ_{er}^l 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. (70) 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. (75) becomes

$$|c_r(t)|^2 = \frac{1}{4}\left|1 - e^{(4i\frac{g^2}{\Delta\omega_t} - \frac{g^2}{\delta^2}\kappa)t}\right|^2. \tag{78}$$

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

$$\mathcal{F} \approx 1 - \frac{4g^2}{\Delta\omega_t^2}\kappa T_g \approx R. \tag{79}$$

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

$$\mathcal{F} \approx R - \frac{\pi \Delta \omega_t}{16 g^2 T_2^*}. \tag{80}$$

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 $|\Phi_t^{\text{out,L}}|$.

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 $|\Phi_l^{\text{out,L}}|$ 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

$$
\begin{aligned}
\Phi_{n,r}^{\text{out,L}}(t) &= \Phi_{n,r}^{\text{in,L}}(t) + \sqrt{\frac{\gamma_{n,r}(t)}{2}} c_r(t) e^{i\theta_r(t)}, \\
&= -\Phi_{n,r}^{\text{out,R}}(t) e^{i\phi_r^n} + \sqrt{\frac{\gamma_{n,r}(t)}{2}} c_r(t) e^{i\theta_r(t)}, \\
&= -\left[\Phi_{n,r}^{\text{in,R}}(t) + \sqrt{\frac{\gamma_{n,r}(t)}{2}} c_r(t) e^{i\theta_r(t)}\right] e^{i\phi_r^n} + \sqrt{\frac{\gamma_{n,r}(t)}{2}} c_r(t) e^{i\theta_r(t)}, \\
&= -\Phi_{n,e}^{\text{out,R}}(t) e^{i(\phi_r^n + \phi_{er}^n)} + \frac{1 - e^{i\phi_r^n}}{2} \sqrt{\gamma_r(t)} c_r(t) e^{i\theta_r(t)}.
\end{aligned}
\tag{81}
$$

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 $\Phi_{n,r}^{\text{out,L}}$ never reaches back the emitting center. In that case, the output field of the emitter simplifies to ($\beta_e^n = 0.5$)

$$
\begin{aligned}
\Phi_{n,e}^{\text{out,R}}(t) &= \Phi_{n,e}^{\text{in,R}}(t) + \sqrt{\frac{\gamma_{n,e}(t)}{2}} c_e(t) e^{i\theta_e(t)}, \\
&= -\Phi_{n,e}^{\text{out,L}}(t) e^{i\phi_e^n} + \sqrt{\frac{\gamma_{n,e}(t)}{2}} c_e(t) e^{i\theta_e(t)}, \\
&= \frac{1 - e^{i\phi_e^n}}{2} \sqrt{\gamma_e(t)} c_e(t) e^{i\theta_e(t)}, \\
&\equiv \frac{1 - e^{i\phi_e^n}}{2} \Phi(t).
\end{aligned}
\tag{82}
$$

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

$$\sqrt{\gamma_r(t)} c_r(t) e^{i\theta_r(t)} = \frac{\sin(\phi_e^t/2)}{\sin(\phi_r^t/2)} \Phi(t) e^{i\phi_L^t/2} \quad \therefore \quad \phi_L^t = \phi_r^t + \phi_e^t + 2\phi_{er}^t, \tag{83}$$

the left-propagating longitudinal signal becomes

$$
\begin{aligned}
r_l = \left|\frac{\Phi_{l,r}^{\text{out,L}}(t)}{\Phi(t) \sin(\phi_e^l/2)}\right|^2 &= \left|1 - \frac{\sin(\phi_e^t/2)}{\sin(\phi_e^l/2)} \frac{\sin(\phi_r^l/2)}{\sin(\phi_r^t/2)} e^{i(\phi_L^t - \phi_L^l)/2}\right|^2, \\
&= 1 + \frac{\sin^2(\phi_e^t/2)}{\sin^2(\phi_e^l/2)} \frac{\sin^2(\phi_r^l/2)}{\sin^2(\phi_r^t/2)} - 2\frac{\sin(\phi_e^t/2)}{\sin(\phi_e^l/2)} \frac{\sin(\phi_r^l/2)}{\sin(\phi_r^t/2)} \cos[(\phi_L^t - \phi_L^l)/2].
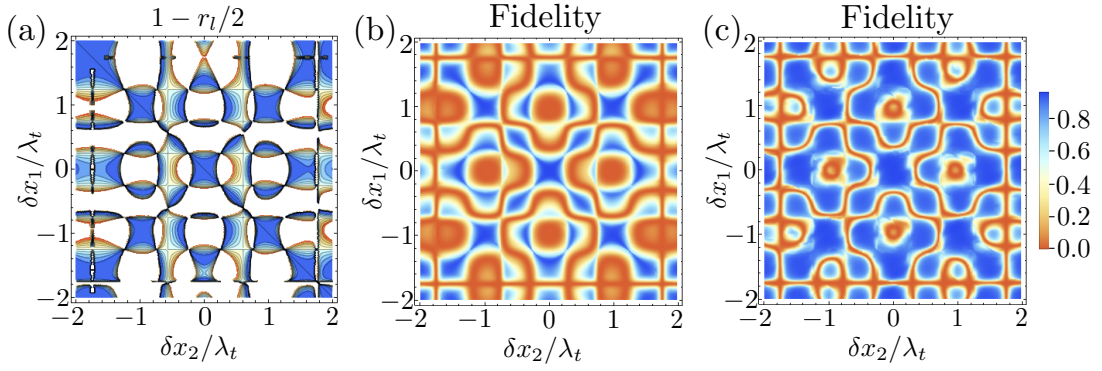\end{aligned}
\tag{84}
$$

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. (84). (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.

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. (84). 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. (84) gives the proper intuition. Already at this level, the fidelity is robust for small variations, as predicted by a small displacement expansion

$$r_l \approx \frac{(k_l - k_t)^2}{4}(\delta x_e^2 + \delta x_r^2)^2. \tag{85}$$

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, B. Pingault, L. Shao, H. A. Atikian, J. Holzgrafe, M. Gundogan, C. Stavrakas, A. Sipahigil, C. Chia, M. J. Burek, M. Zhang, L. Wu, J. L. Pacheco, J. Abraham, E. Bielejec, M. D. Lukin, M. Atature, M. Loncar, Strain engineering of the silicon vacancy center in diamond, arXiv:1801.09833.

[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).

[8] 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).

[9] P. C. Maurer, J. R. Maze, P. L. Stanwix, L. Jiang, A. V. Gorshkov, A. A. Zibrov, B. Harke, J. S. Hodges, A. S. Zibrov, A. Yacoby, D. Twitchen, S. W. Hell, R. L. Walsworth and M. D. Lukin, Far-field optical imaging and manipulation of individual spins with nanoscale resolution, Nature Physics **6**, 912-918 (2010).

[10] R. Li, L. Petit, D.P. Franke, J.P. Dehollain, J. Helsen, M. Steudtner, N.K. Thomas, Z.R. Yoscovits, K.J. Singh, S. Wehner, L.M.K. Vandersypen, J.S. Clarke, and M. Veldhorst, A Crossbar Network for Silicon Quantum Dot Qubits, arXiv:1711.03807.