

Supplemental information for  
**Coherent Atom-Phonon Interaction through Mode Field  
Coupling in Hybrid Optomechanical Systems**

Michele Cotrufo\* and Andrea Fiore

*Department of Applied Physics, Eindhoven University of Technology,  
5600 MB Eindhoven, The Netherlands*

Ewold Verhagen

*Center for Nanophotonics, AMOLF,  
Science Park 104, 1098 XG Amsterdam, The Netherlands*

---

\* Electronic address: m.cotrufo@tue.nl

## CONTENTS

I. Density matrix calculations, numerical setup	2
II. Derivation of the Mode Field Coupling for the two-cavity system	3
III. Derivation of the Mode Field Coupling for the three-cavity system	5
IV. Validity of the linear approximation in the derivation of the Mode Field Coupling	8
V. Emitter decay rate in the three-cavity system	9
VI. Additional radiative losses introduced by the supermodes $\hat{a}_+$ and $\hat{a}_-$	10
VII. Selective pumping of the optical mode of interest in the three-cavity system	10
VIII. Three-cavity system Hamiltonian at large driving	11
IX. Influence of cavity population and dephasing on the fidelity of the one-phonon state creation	11
X. Creation of other nonclassical mechanical states	12
X.1. One-cavity model with MFC	12
X.2. Three-cavity model	14
XI. Role of the other two mechanical supermodes	15
XII. Design of the proposed structure	17
XIII. Tripartite interaction in a Fabry-Perot cavity	18
References	19

## I. DENSITY MATRIX CALCULATIONS, NUMERICAL SETUP

The non-Hermitian evolution of the different systems considered in this work has been calculated with a Master equation approach, in which the different dissipative channels are described by proper Lindblad terms. For a given Hamiltonian  $\hat{H}$ , we calculate the temporal evolution of the density matrix  $\rho$  through the equation

$$\begin{aligned} \dot{\rho} = & -\frac{i}{\hbar} [\hat{H}, \rho] + \sum_i \frac{\kappa_i}{2} \left( 2\hat{a}_i \rho \hat{a}_i^\dagger - \{ \hat{a}_i^\dagger \hat{a}_i, \rho \} \right) + \frac{\Gamma_M (n_{\text{th}} + 1)}{2} \left( 2\hat{b} \rho \hat{b}^\dagger - \{ \hat{b}^\dagger \hat{b}, \rho \} \right) + \\ & + \frac{\Gamma_M n_{\text{th}}}{2} \left( 2\hat{b}^\dagger \rho \hat{b} - \{ \hat{b} \hat{b}^\dagger, \rho \} \right) + \frac{\Gamma}{2} (2\hat{\sigma}_- \rho \hat{\sigma}_+ - \{ \hat{\sigma}_+ \hat{\sigma}_-, \rho \}) + \frac{\Gamma^*}{2} (\hat{\sigma}_z \rho \hat{\sigma}_z - \rho) \end{aligned} \quad (\text{S1})$$

where the sum refers to all the optical cavities considered,  $\kappa_i$  is decay rate of the  $i$ -th cavity,  $\Gamma_M$  is the decay rate of the mechanical mode,  $n_{\text{th}}$  is the average phonon number of the external bath, and  $\Gamma$  and  $\Gamma^*$  are the decay rate and pure dephasing rate of the emitter, respectively. The brackets  $[\cdot]$  and  $\{\cdot\}$  indicate commutation and anti-commutation of the operators, respectively. The master equation has been solved numerically with the opensource Python framework QuTIP [1]. For the numerical calculations, the dimensions

of the Fock spaces of the optical cavities and mechanical resonator need to be truncated. For the calculations of the emitter-phonon swapping, the Fock spaces of the mechanical oscillator and optical cavities all have dimensions of 2. We verified that no appreciable numerical differences arise for larger Fock spaces dimensions. For the cooling calculations, the Fock space of the mechanical resonator has a dimension of 15, to ensure that a thermal state with  $n_{\text{th}} = 4$  can be properly described.

Both in the in main text and in this SI we frequently use the concept of fidelity to quantify the closeness of a created state and a target one. The fidelity is defined as  $F(\hat{\rho}_A, \hat{\rho}_B) \equiv \text{tr} \left( \sqrt{(\hat{\rho}_A)^{1/2} \hat{\rho}_B (\hat{\rho}_A)^{1/2}} \right)$  [2], where  $\hat{\rho}_A$  and  $\hat{\rho}_B$  are the density matrices of the states to be compared and “tr” denotes the trace.

## II. DERIVATION OF THE MODE FIELD COUPLING FOR THE TWO-CAVITY SYSTEM

The general model describing mode field coupling in a two-cavity system is schematically depicted in fig. S1. We consider two identical optomechanical systems, denoted left (L) and right (R), each composed of an optical cavity with frequency  $\omega_c$  and a mechanical resonator with frequency  $\Omega_M$ . In each system, the cavity and the resonator are dispersively coupled at a rate  $g_0$ . We describe the fields in the two optical cavities with the annihilation operators  $\hat{a}_L$  and  $\hat{a}_R$ , and the two mechanical resonators with annihilation operators  $\hat{b}_L$  and  $\hat{b}_R$ . The two optical cavities are coupled with a rate  $J$ , while the two mechanical resonators are coupled with a rate  $J_M$ . Finally, a two-level emitter is placed in the left cavity, and interacts with the field  $\hat{a}_L$  with a coupling rate  $g$ .

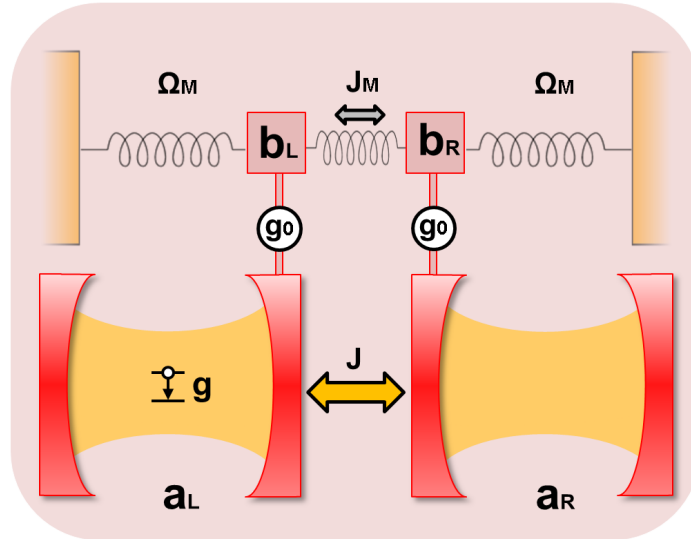


FIG. S1. Schematic of the two-cavity system. Two identical optical cavities (each denoted by a couple of red mirrors) interact with each other at a rate  $J$ . Each cavity is dispersively coupled to a separate mechanical resonator with an optomechanical coupling rate  $g_0$ . The resonators are identical and have frequency  $\Omega_M$ . The two mechanical resonators are additionally coupled to each other at a rate  $J_M$ . An emitter is placed in the left optical cavity and interacts with one of its optical modes with a coupling rate  $g$ .

The full Hamiltonian reads

$$\begin{aligned} \hat{H} = & \left[ \omega_c - g_0 \left( \hat{b}_L^\dagger + \hat{b}_L \right) \right] \hat{a}_L^\dagger \hat{a}_L + \left[ \omega_c - g_0 \left( \hat{b}_R^\dagger + \hat{b}_R \right) \right] \hat{a}_R^\dagger \hat{a}_R + \Omega_M \left( \hat{b}_L^\dagger \hat{b}_L + \hat{b}_R^\dagger \hat{b}_R \right) + \frac{\omega_A}{2} \hat{\sigma}_z + \\ & + J \left( \hat{a}_R^\dagger \hat{a}_L + h.c. \right) + J_M \left( \hat{b}_R^\dagger \hat{b}_L + h.c. \right) + g \left( \hat{\sigma}_+ \hat{a}_L + h.c. \right). \end{aligned} \quad (S2)$$

We now introduce mechanical supermodes  $\hat{b}_\pm = \frac{1}{\sqrt{2}} (\hat{b}_L \pm \hat{b}_R)$ . Substituting these in the Hamiltonian, we obtain

$$\begin{aligned} \hat{H} = & \left[ \omega_c - \frac{g_0}{\sqrt{2}} (\hat{x}_+ + \hat{x}_-) \right] \hat{a}_L^\dagger \hat{a}_L + \left[ \omega_c - \frac{g_0}{\sqrt{2}} (\hat{x}_+ - \hat{x}_-) \right] \hat{a}_R^\dagger \hat{a}_R + \\ & + (\Omega_M + J_M) \hat{b}_+^\dagger \hat{b}_+ + (\Omega_M - J_M) \hat{b}_-^\dagger \hat{b}_- + \frac{\omega_A}{2} \hat{\sigma}_z + \\ & + J \left( \hat{a}_R^\dagger \hat{a}_L + h.c. \right) + g \left( \hat{\sigma}_+ \hat{a}_L + h.c. \right), \end{aligned} \quad (S3)$$

where we have defined the dimensionless position operators of the mechanical supermodes  $\hat{x}_\pm = \hat{b}_\pm^\dagger + \hat{b}_\pm$ . The mechanical mode  $\hat{b}_-$  interacts dispersively with the modes  $\hat{a}_R$  and  $\hat{a}_L$  with equal and opposite coupling rate  $\pm g_0/\sqrt{2}$ . In the following, we will neglect the presence of the other mechanical supermode,  $\hat{b}_+$ . For simplicity, we replace  $\hat{b}_- \rightarrow \hat{b}$  and  $(\Omega_M - J_M) \rightarrow \Omega_M$ . We moreover define the operator  $\hat{\Delta} = \frac{g_0}{\sqrt{2}} (\hat{b}^\dagger + \hat{b})$ . The new Hamiltonian reads

$$\begin{aligned} \hat{H} = & -\hat{\Delta} \hat{a}_L^\dagger \hat{a}_L + \hat{\Delta} \hat{a}_R^\dagger \hat{a}_R + \Omega_M \hat{b}^\dagger \hat{b} + \frac{\omega_A - \omega_c}{2} \hat{\sigma}_z + \\ & + J \left( \hat{a}_R^\dagger \hat{a}_L + h.c. \right) + g \left( \hat{\sigma}_+ \hat{a}_L + h.c. \right), \end{aligned} \quad (S4)$$

where we also performed a unitary transformation  $\hat{H} \rightarrow \hat{U}(t) \hat{H} \hat{U}^\dagger(t) - i \hat{U}(t) \frac{\partial \hat{U}(t)^\dagger}{\partial t}$ , with  $\hat{U}(t) = \exp \left[ -i \omega_c t \left( \hat{a}_L^\dagger \hat{a}_L + \hat{a}_R^\dagger \hat{a}_R + \hat{\sigma}_+ \hat{\sigma}_- \right) \right]$ . Assuming a quasi-static approximation for  $\hat{\Delta}$ , valid in the limit  $J \gg \Omega_M$ , we can diagonalize the optical part of the Hamiltonian [3]. The optical supermodes are defined by

$$\hat{a}_+ = \alpha \hat{a}_L + \beta \hat{a}_R, \quad \hat{a}_- = \beta \hat{a}_L - \alpha \hat{a}_R, \quad (S5)$$

where  $\alpha$  and  $\beta$  are operators defined by

$$\alpha = \frac{\sqrt{\hat{\Delta}^2 + J^2} + \hat{\Delta}}{\sqrt{\left( \sqrt{\hat{\Delta}^2 + J^2} + \hat{\Delta} \right)^2 + J^2}}, \quad \beta = \frac{J}{\sqrt{\left( \sqrt{\hat{\Delta}^2 + J^2} + \hat{\Delta} \right)^2 + J^2}}. \quad (S6)$$

In the supermode basis, the Hamiltonian is

$$\begin{aligned} \hat{H} = & \sqrt{J^2 + \hat{\Delta}^2} \hat{a}_+^\dagger \hat{a}_+ - \sqrt{J^2 + \hat{\Delta}^2} \hat{a}_-^\dagger \hat{a}_- + \frac{\omega_A - \omega_c}{2} \hat{\sigma}_z + \Omega_M \hat{b}^\dagger \hat{b} + \\ & + g \alpha \left( \hat{\sigma}_+ \hat{a}_+ + h.c. \right) + g \beta \left( \hat{\sigma}_+ \hat{a}_- + h.c. \right). \end{aligned} \quad (S7)$$

Up to the first order in  $\hat{\Delta}/J$ , the coefficients  $\alpha$  and  $\beta$  read

$$\alpha = \frac{1}{\sqrt{2}} \left( 1 + \frac{\hat{\Delta}}{2J} \right) + \mathcal{O} \left( \frac{\hat{\Delta}}{J} \right)^2, \quad \beta = \frac{1}{\sqrt{2}} \left( 1 - \frac{\hat{\Delta}}{2J} \right) + \mathcal{O} \left( \frac{\hat{\Delta}}{J} \right)^2. \quad (\text{S8})$$

By inserting these expansions in the Hamiltonian in eq. S7, and expressing again  $\hat{\Delta}$  as a function of  $\hat{b}$  and  $\hat{b}^\dagger$ , we get

$$\begin{aligned} \hat{H} = & \sqrt{J^2 + \hat{\Delta}^2} \hat{a}_+^\dagger \hat{a}_+ - \sqrt{J^2 + \hat{\Delta}^2} \hat{a}_-^\dagger \hat{a}_- + \frac{\omega_A - \omega_c}{2} \hat{\sigma}_z + \Omega_M \hat{b}^\dagger \hat{b} + \\ & + \frac{g}{\sqrt{2}} (\hat{\sigma}_+ \hat{a}_+ + h.c.) + \frac{g}{\sqrt{2}} (\hat{\sigma}_+ \hat{a}_- + h.c.) + \\ & + \frac{gg_0}{4J} (\hat{b}^\dagger + \hat{b}) (\hat{\sigma}_+ \hat{a}_+ + h.c.) - \frac{gg_0}{4J} (\hat{b}^\dagger + \hat{b}) (\hat{\sigma}_+ \hat{a}_- + h.c.). \end{aligned} \quad (\text{S9})$$

The second row describes the Rabi interaction of the emitter with the two optical supermodes, with a coupling rate  $g/\sqrt{2}$ . As mentioned in the main text, in the two-cavity system this interaction is unavoidable and is due to the fact that both optical supermodes have nonzero field in both cavities for all finite detunings. The third row describes the tripartite interaction between the emitter, the phonon and the optical supermodes.

### III. DERIVATION OF THE MODE FIELD COUPLING FOR THE THREE-CAVITY SYSTEM

The model for the three-cavity case is schematically depicted in fig. S2. The three identical optical cavities, denoted left (L), target (T) and right (R) are dispersively coupled at a rate  $g_0$  with a separate mechanical resonator, in similar fashion as in the previous section. The three mechanical resonators are assumed identical and with frequency  $\Omega_M$ . We describe the three optical cavities with the annihilation operators  $\hat{a}_L$ ,  $\hat{a}_T$  and  $\hat{a}_R$ , and the three resonators with annihilation operators  $\hat{b}_L$ ,  $\hat{b}_T$  and  $\hat{b}_R$ . The optical cavity T is coupled to the cavities L and R with a rate  $J$ , while the mechanical resonator  $\hat{b}_T$  is coupled with a rate  $J_M$  to the resonators  $\hat{b}_R$  and  $\hat{b}_L$ . Finally, a two-level emitter is placed in the target cavity, and interacts with the field  $\hat{a}_T$  with a coupling rate  $g$ .

The full Hamiltonian reads

$$\begin{aligned} \hat{H} = & \left[ \omega_c - g_0 (\hat{b}_L^\dagger + \hat{b}_L) \right] \hat{a}_L^\dagger \hat{a}_L + \left[ \omega_c - g_0 (\hat{b}_T^\dagger + \hat{b}_T) \right] \hat{a}_T^\dagger \hat{a}_T + \left[ \omega_c - g_0 (\hat{b}_R^\dagger + \hat{b}_R) \right] \hat{a}_R^\dagger \hat{a}_R + \\ & + \Omega_M (\hat{b}_L^\dagger \hat{b}_L + \hat{b}_T^\dagger \hat{b}_T + \hat{b}_R^\dagger \hat{b}_R) + \frac{\omega_A}{2} \hat{\sigma}_z + \\ & + J \left[ \hat{a}_T^\dagger (\hat{a}_R + \hat{a}_L) + h.c. \right] + J_M \left[ \hat{b}_T^\dagger (\hat{b}_R + \hat{b}_L) + h.c. \right] + g (\hat{\sigma}_+ \hat{a}_T + h.c.). \end{aligned} \quad (\text{S10})$$

The mechanical part of this Hamiltonian can be diagonalized by introducing the modes

$$\begin{aligned} \hat{b}_0 &= \frac{1}{\sqrt{2}} (\hat{b}_L - \hat{b}_R), \\ \hat{b}_+ &= \frac{1}{\sqrt{2}} \hat{b}_T + \frac{1}{2} (\hat{b}_R + \hat{b}_L), \\ \hat{b}_- &= -\frac{1}{\sqrt{2}} \hat{b}_T + \frac{1}{2} (\hat{b}_R + \hat{b}_L). \end{aligned} \quad (\text{S11})$$

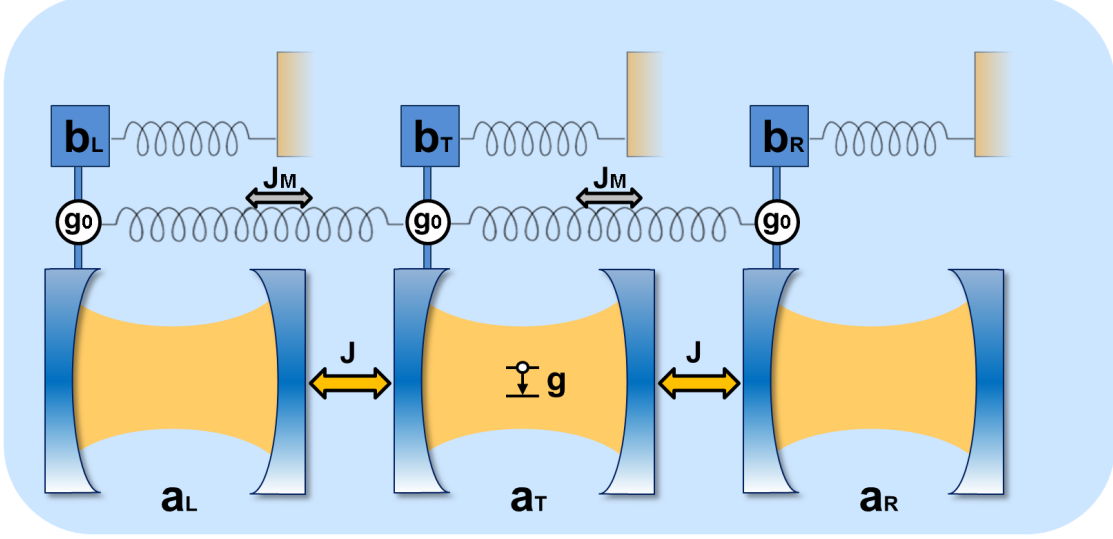


FIG. S2. Schematic of the three-cavity system. Three identical optical cavities (each denoted by a couple of blue mirrors) are arranged such that the central one interacts with the two lateral ones with a rate  $J$ . Each cavity is dispersively coupled to a separate resonator with an optomechanical coupling rate  $g_0$ . The resonators are identical and have frequency  $\Omega_M$ . The central mechanical resonator is additionally coupled to the other two mechanical resonators with a rate  $J_M$ . An emitter is placed in the central optical cavity and interacts with its optical mode with a coupling rate  $g$ .

The transformed Hamiltonian reads

$$\begin{aligned} \hat{H} = & \omega_c \left( \hat{a}_L^\dagger \hat{a}_L + \hat{a}_T^\dagger \hat{a}_T + \hat{a}_R^\dagger \hat{a}_R \right) + \Omega_M \hat{b}_0^\dagger \hat{b}_0 + \left( \Omega_M + \sqrt{2}J_M \right) \hat{b}_+^\dagger \hat{b}_+ + \left( \Omega_M - \sqrt{2}J_M \right) \hat{b}_-^\dagger \hat{b}_- + \frac{\omega_A}{2} \hat{\sigma}_z + \\ & + \frac{g_0}{\sqrt{2}} \hat{x}_0 \left( \hat{a}_R^\dagger \hat{a}_R - \hat{a}_L^\dagger \hat{a}_L \right) + \frac{g_0}{2} (\hat{x}_+ - \hat{x}_-) \left( \hat{a}_L^\dagger \hat{a}_L + \hat{a}_R^\dagger \hat{a}_R \right) - \frac{g_0}{2} (\hat{x}_+ - \hat{x}_-) \hat{a}_T^\dagger \hat{a}_T + \\ & + J \left[ \hat{a}_T^\dagger (\hat{a}_R + \hat{a}_L) + h.c. \right] + g (\hat{\sigma}_+ \hat{a}_T + h.c.), \end{aligned} \quad (\text{S12})$$

where we have defined the dimensionless position operators of the mechanical supermodes  $\hat{x}_\pm = \hat{b}_\pm^\dagger + \hat{b}_\pm$  and  $\hat{x}_0 = \hat{b}_0^\dagger + \hat{b}_0$ . We now focus only on the mechanical mode  $\hat{b}_0$  with the assumption that, under the resonant condition  $\omega_A = \omega_c + \Omega_M$ , the terms involving the other mechanical modes are negligible (this assumption is numerically verified in section XI). For simplicity, we redefine  $\hat{b}_0 \rightarrow \hat{b}$ . Moreover, we perform a unitary transformation

$$\hat{H} \rightarrow \hat{U}(t) \hat{H} \hat{U}^\dagger(t) - i \hat{U}(t) \frac{\partial \hat{U}(t)^\dagger}{\partial t}, \quad \text{with } \hat{U}(t) = \exp \left[ -i \omega_c t \left( \hat{a}_L^\dagger \hat{a}_L + \hat{a}_T^\dagger \hat{a}_T + \hat{a}_R^\dagger \hat{a}_R + \hat{\sigma}_+ \hat{\sigma}_- \right) \right].$$

The new Hamiltonian reads

$$\begin{aligned} \hat{H} = & -\hat{\Delta} \hat{a}_L^\dagger \hat{a}_L + \hat{\Delta} \hat{a}_R^\dagger \hat{a}_R + \Omega_M \hat{b}^\dagger \hat{b} + \frac{\omega_A - \omega_c}{2} \hat{\sigma}_z + \\ & + J \left[ \hat{a}_T^\dagger (\hat{a}_R + \hat{a}_L) + h.c. \right] + g (\hat{\sigma}_+ \hat{a}_T + h.c.), \end{aligned} \quad (\text{S13})$$

where we defined  $\hat{\Delta} = g_0 \hat{x}_0 / \sqrt{2}$ . We note that the selected mechanical mode has equal and opposite dispersive coupling with the optical cavities L and R, at a rate  $\pm g_0 / \sqrt{2}$ , while it does not affect the optical cavity T. We again assume a quasi-static approximation for  $\hat{\Delta}$ , valid in the limit  $J \gg \Omega_M$ , to diagonalize the optical part of the Hamiltonian [3]. The three

optical supermodes are defined in this case by

$$\begin{aligned}\hat{a}_0 &= -\epsilon\hat{a}_L + \beta\hat{a}_T + \epsilon\hat{a}_R, \\ \hat{a}_+ &= \eta\hat{a}_L - \epsilon\hat{a}_T + \mu\hat{a}_R, \\ \hat{a}_- &= \mu\hat{a}_L + \epsilon\hat{a}_T + \eta\hat{a}_R,\end{aligned}\tag{S14}$$

where  $\epsilon$ ,  $\beta$ ,  $\mu$  and  $\eta$  are operators. In particular,

$$\epsilon = \frac{1}{\sqrt{2 + \left(\hat{\Delta}/J\right)^2}}, \quad \beta = \frac{\hat{\Delta}/J}{\sqrt{2 + \left(\hat{\Delta}/J\right)^2}}.\tag{S15}$$

The functions  $\eta$  and  $\mu$  satisfy  $\eta(\hat{\Delta}, J) = \mu(-\hat{\Delta}, J)$ . Their expressions are more complicated and not reported here, as they are not needed in the following. The Hamiltonian in the supermode basis reads

$$\begin{aligned}\hat{H} &= \sqrt{2J^2 + \hat{\Delta}^2} \left( \hat{a}_+^\dagger \hat{a}_+ - \hat{a}_-^\dagger \hat{a}_- \right) + \Omega_M \hat{b}^\dagger \hat{b} + \frac{\omega_A - \omega_c}{2} \hat{\sigma}_z + \\ &+ g\beta (\hat{\sigma}_+ \hat{a}_0 + h.c.) + g\epsilon [\hat{\sigma}_+ (\hat{a}_+ - \hat{a}_-) + h.c.].\end{aligned}\tag{S16}$$

Up to the first order in  $\hat{\Delta}/J$ , the functions  $\beta$  and  $\epsilon$  are

$$\beta = \frac{1}{\sqrt{2}} \frac{\hat{\Delta}}{J} + \mathcal{O}\left(\frac{\hat{\Delta}}{J}\right)^3, \quad \epsilon = \frac{1}{\sqrt{2}} + \mathcal{O}\left(\frac{\hat{\Delta}}{J}\right)^2.\tag{S17}$$

By inserting these expansions in the Hamiltonian in eq. S16, and expressing again  $\hat{\Delta}$  as a function of  $\hat{b}$  and  $\hat{b}^\dagger$ , we get

$$\begin{aligned}\hat{H} &= \sqrt{2J^2 + \hat{\Delta}^2} \left( \hat{a}_+^\dagger \hat{a}_+ - \hat{a}_-^\dagger \hat{a}_- \right) + \Omega_M \hat{b}^\dagger \hat{b} + \frac{\omega_A - \omega_c}{2} \hat{\sigma}_z + \\ &+ \frac{gg_0}{2J} \left( \hat{b}^\dagger + \hat{b} \right) (\hat{\sigma}_+ \hat{a}_0 + h.c.) + \frac{g}{\sqrt{2}} [\hat{\sigma}_+ (\hat{a}_+ - \hat{a}_-) + h.c.].\end{aligned}\tag{S18}$$

The first term of the second row describes the mode field coupling between the emitter, the phonon and the mode  $\hat{a}_0$ , with a coupling rate  $\gamma = \frac{gg_0}{2J}$ . The second term of the second row describes a Rabi interaction between the emitter and the *other* two optical supermodes. As confirmed by the numerical simulations based on the full Hamiltonian in eq. S13 (see main text), the effect of these terms is negligible when the spectral separation between the supermodes is much larger than the mechanical frequency ( $\sqrt{2}J \gg \Omega_M$ ). However, for finite optical linewidths ( $\kappa$ ), the presence of the modes  $\hat{a}_\pm$  will introduce additional decay channels for the QE, as explained below.

To verify the validity of the effective Hamiltonian derived in eq. S18, and in particular the presence of the predicted tripartite QE-phonon-photon interaction, we numerically calculated the evolution of the system dictated by the Hamiltonian in eq. S13, *i.e.*, without introducing the optical supermodes. We start with the QE excited, we use the parameters mentioned in the main text,  $\{\omega_c, \Omega_M, g, g_0, J\} = 2\pi \cdot \{4.7 \cdot 10^5, 14, 20, 0.004, 360\}$  GHz, and we set the resonant condition  $\omega_A = \omega_c + \Omega_M$ . As shown in fig. S3a, the lossless evolution of the system agrees perfectly with that of the MFC Hamiltonian (fig. 1c of the main text): the system oscillates between the state in which only the QE is excited and the state in which

the QE is in the ground state and one photon and one phonon are present. The oscillation period verifies the predicted value of  $T = \pi/\gamma = 2\pi J/(gg_0) = 4.5 \mu\text{s}$ . This confirms that, for realistic choices of parameters, a purely tripartite interaction is obtained in the three-cavity system, as predicted by the effective Hamiltonian in eq. S18.

Even in presence of losses, nonclassical states can be obtained. In the main text we have shown how the losses can be overcome by strongly pumping the mode  $\hat{a}_0$  (fig. 3a of the main text), and the one-phonon state can be created with high fidelity,  $F=0.86$ . The Wigner map of the mechanical state created in fig. 3a is shown in fig. S3b. The nonclassical nature of the state is evident from the negativity of the Wigner map.

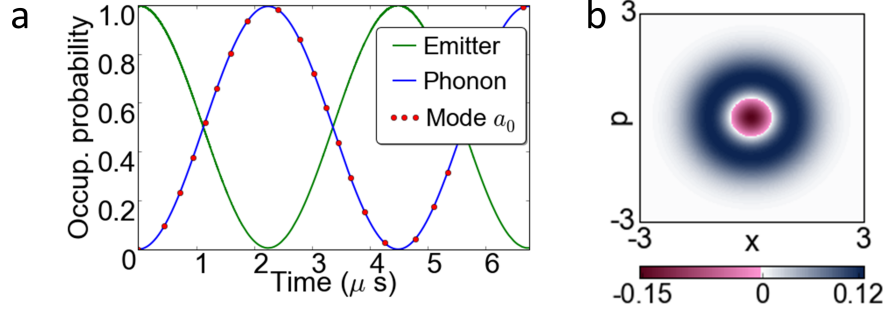


FIG. S3. (a) Lossless evolution dictated by the Hamiltonian in eq. S13, starting from an excited QE and with a red-detuned cavity ( $\omega_A = \omega_c + \Omega_M$ ). The parameters are  $\{\omega_c, \Omega_M, g, g_0, J\} = 2\pi \cdot \{4.7 \cdot 10^5, 14, 20, 0.004, 360\}$  GHz. (b) Wigner map of the mechanical state created in fig. 3a of the main paper.

#### IV. VALIDITY OF THE LINEAR APPROXIMATION IN THE DERIVATION OF THE MODE FIELD COUPLING

In deriving the Hamiltonian of both the two-cavity and three-cavity systems we neglected terms of second or higher order in  $\hat{\Delta}/J$  (eqs. S8 and S17). This approximation sets, in principle, a limit on the mechanical nonclassical state which can be created via the MFC, because for a generic state  $|\Psi\rangle$  the expectation value of the operator  $\hat{\Delta}$  may be not negligible with respect to  $J$ . However, as we show here, this limit is extremely large. We can consider a generic state of the mechanical resonator,  $|\Psi\rangle = \sum_{n=0}^{\infty} C_n |n\rangle$ , where the  $C_n$  are arbitrary coefficients and  $|n\rangle$  are the Fock states. Since the expectation value of the operator  $\hat{\Delta} = g_0/\sqrt{2}(\hat{b} + \hat{b}^\dagger)$  is zero in any Fock state, we instead consider the quantity  $\sqrt{\langle \hat{\Delta}^2 \rangle}$ . Requiring that  $\sqrt{\langle \hat{\Delta}^2 \rangle} \ll J$  leads to

$$\sum_{n=0}^{\infty} \left[ (C_{n+2}^* C_n + C_{n+2} C_n^*) \sqrt{(n+1)(n+2)} + |C_n|^2 (2n+1) \right] \ll \frac{2J^2}{g_0^2}. \quad (\text{S19})$$

For a Fock state (*i.e.*,  $C_n = 1, C_{m \neq n} = 0$ ), this inequality leads to  $n \ll (J/g_0)^2$ . With the parameters used in the main text, this translates into  $n \ll 10^{10}$ , which ensures us that the MFC allows generating nonclassical mechanical states in extremely large Fock spaces. Essentially similar results hold if one consider a superposition of a finite number of Fock states. A similar analysis can be carried out for a coherent state  $|\Psi\rangle = |\alpha\rangle$ , which leads to the limit  $|\alpha| \ll J/g_0 \approx 10^5$ .



## V. EMITTER DECAY RATE IN THE THREE-CAVITY SYSTEM

As mentioned in the main text, for the calculation in fig. 3b we assume a conservative value of  $\Gamma/2\pi = 0.05$  GHz for the rate of the emitter decay into the leaky modes (*i.e.*, all the optical modes different from the three supermodes of the system). This value corresponds to the one measured for an NV center placed in a single defect cavity (similar to the one used here to realize a three-cavity system) and out of resonance with the optical mode [4]. However, significantly smaller radiative decay rates could in principle occur in the proposed structure, as we explain in the following. To evaluate the expected decay rate, we perform finite element method calculations (COMSOL) of the power emitted by an electric dipole (fig. S4) placed in either the target (blue line) or the left cavity (green line) of the structure considered in the main text. We then normalize the calculated powers by the power emitted by the same dipole in bulk diamond. In this way we obtain a prediction for the radiative decay rate of a QE (with the same position and polarization of the dipole) normalized to the radiative decay rate of the same QE in bulk. In agreement with the spatial patterns of the optical supermodes (see also sec. XII of this S.I.), a dipole placed in the lateral cavity can feed all the three optical supermodes, while one that is located in the target cavity does not emit into the mode  $\hat{a}_0$ . In particular, the calculation shows that the radiative decay rate of a QE placed in the cavity T is reduced by almost a factor  $10^3$  with respect to the decay rate in bulk. As the measured lifetime of an NV center in bulk is about 12 ns [5], a radiative decay rate of the order of  $\Gamma/2\pi = 0.1$  MHz is in principle expected for the same emitter placed in the target cavity. We also note that the emission of a dipole in the target cavity is dominated by the coupling with the nonresonant coupled modes for the chosen set of parameters (see sec. VI), while the emission into leaky modes is expected to be even smaller. We therefore conclude that the assumed value of 50 MHz is indeed a conservative, upper estimate of the QE decay rate.

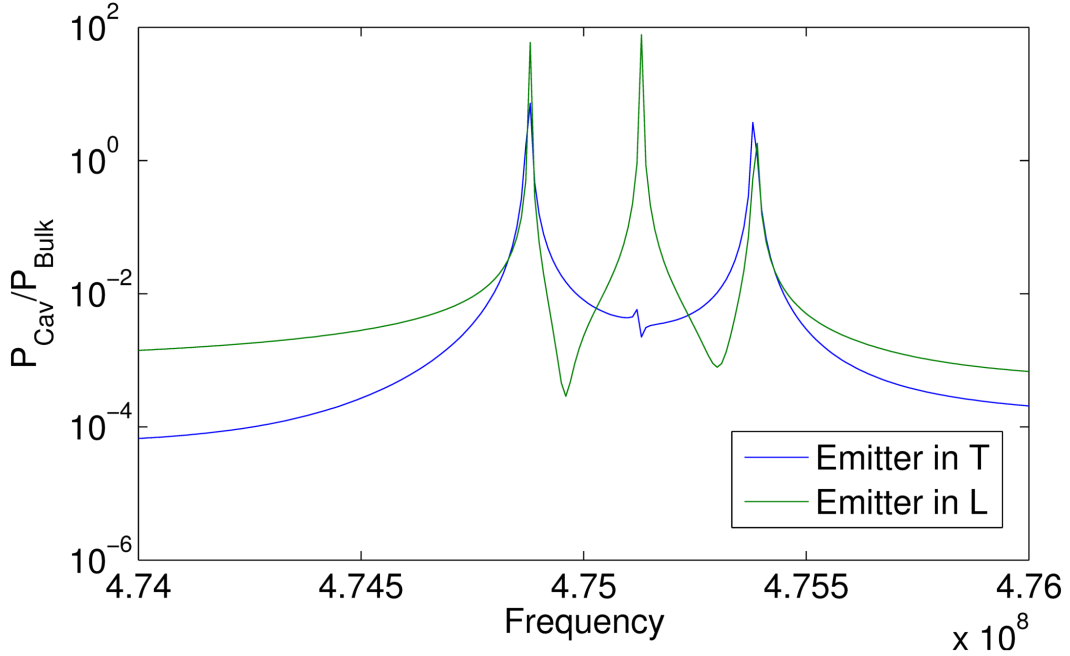


FIG. S4. Calculated power emitted by a dipole placed in the target (blue line) and left (green line) cavity of the three-cavity diamond structure discussed in the main text and in sec. XII of this SI. The powers ( $P_{\text{Cav}}$ ) are normalized by the power emitted in bulk diamond ( $P_{\text{Bulk}}$ ).

## VI. ADDITIONAL RADIATIVE LOSSES INTRODUCED BY THE SUPERMODES $\hat{a}_+$ AND $\hat{a}_-$

As shown in the main text, for a proper parameter choice the Hamiltonian of the three-cavity system provides a tripartite interaction between the supermode  $\hat{a}_0$ , the QE and the mechanical mode. The interaction between the QE and the modes  $\hat{a}_\pm$  is suppressed for large supermode separation ( $\sqrt{2}J \gg \Omega_M$ ). However, these optical modes are still coupled with the emitter (since their fields at the emitter position is not zero), as shown in fig. S4. For large optical losses  $\kappa > g$ , the supermodes  $\hat{a}_\pm$  are weakly coupled to the QE and therefore they introduce additional decay channels (denoted  $\Gamma^{(\pm)}$ ). These additional decay channels become relevant when the other radiative decay channel ( $\Gamma$ ) is small and the optical linewidth is significant. This effect is important in understanding some features related to the fidelity of the created nonclassical states and to the ground-state cooling efficiency (see figs. 3(b,d) and accompanying discussion in the main text).  $\Gamma^{(\pm)}$  can be calculated as follows. If an optical mode with losses  $\kappa$  is resonant with a QE and coupled with it at a rate  $g' < \kappa$ , the QE decays into the optical mode at a rate  $4g'^2/\kappa$ . In the presence of a large spectral detuning  $\delta \gg \kappa$ , the decay rate is modified into  $\frac{4g'^2}{\kappa} \cdot \frac{\kappa^2/4}{\delta^2 + \kappa^2/4} \approx \frac{g'^2\kappa}{\delta^2}$  [6]. In our case,  $\delta = \sqrt{2}J$  and we replace  $g' = g/\sqrt{2}$ , consistently with the notation in the main text where  $g$  is the coupling between the QE and the uncoupled cavity. Therefore,

$$\Gamma^{(\pm)} = \frac{g^2\kappa}{4J^2}. \quad (\text{S20})$$

With respect to the QE-phonon swapping this additional decay channel is negligible when  $\Gamma^{(\pm)} \ll \gamma\sqrt{n_{\text{cav}}} \Rightarrow \kappa \ll 2Jg_0\sqrt{n_{\text{cav}}}/g$ , which corresponds to the vertical dashed-dotted line in fig. 3c of the main text.

## VII. SELECTIVE PUMPING OF THE OPTICAL MODE OF INTEREST IN THE THREE-CAVITY SYSTEM

In the three-cavity system the mode field coupling is due to a particular optical supermode ( $\hat{a}_0$ ), which, when the system is not perturbed by the mechanical displacement, features zero electric field in the central cavity and equal and opposite field amplitudes in the lateral cavities (see the first of eqs. S14 for  $\Delta = 0$ ). In order to enhance the mode field coupling it is therefore necessary to selectively pump this mode, without feeding the other two optical supermodes. This is possible by pumping the two lateral cavities at a frequency  $\omega_c$  with equal field amplitude  $\mathcal{E}$  and opposite phases: excitation of modes  $\hat{a}_\pm$  is then symmetry-forbidden. Focusing only on the optical part of the Hamiltonian, we have

$$\hat{H}_{\text{opt}} = J \left[ \hat{a}_T^\dagger (\hat{a}_R + \hat{a}_L) + h.c. \right] + \mathcal{E} \left( \hat{a}_R^\dagger - \hat{a}_L^\dagger + h.c. \right), \quad (\text{S21})$$

where we switched to a frame rotating at frequency  $\omega_c$ . By writing the equations of motion for the three cavity field amplitudes, and assuming equal cavity losses  $\kappa$ , we can calculate the steady state population under this pumping scheme, namely  $\bar{a}_R = -\bar{a}_L = 2\mathcal{E}/\kappa$  and  $\bar{a}_T = 0$ . By comparing these results with eqs. S14 for  $\Delta = 0$ , we see that only the mode  $\hat{a}_0$  is fed, while the population of the other two modes remains strictly zero. Interestingly, in the limit of validity of the coupled mode theory, this result (*i.e.* the zero population of the supermodes  $\hat{a}_+$  and  $\hat{a}_-$ ) is independent of  $\kappa$ .

### VIII. THREE-CAVITY SYSTEM HAMILTONIAN AT LARGE DRIVING

In the previous section we explained how the mode  $\hat{a}_0$  can be selectively pumped. However, solving a full quantum model (eq. S13) with a large pumping of the optical cavities is computationally extremely challenging, because of the large Fock space dimensions required. We therefore solve a transformed version of the Hamiltonian in eq. S13, obtained through a displacement of the cavity operators. That is, we replace the operators by the sum of a steady-state amplitude and a fluctuating operator, *i.e.*  $\hat{a}_i = \bar{a}_i + \delta\hat{a}_i$ , where  $i = L, R, T$ . The transformed Hamiltonian reads

$$\begin{aligned} \hat{H}_{lin} = & -\hat{\Delta}\delta\hat{a}_L^\dagger\delta\hat{a}_L + \hat{\Delta}\delta\hat{a}_R^\dagger\delta\hat{a}_R + \Omega_M\hat{b}^\dagger\hat{b} + \frac{\omega_A - \omega_c}{2}\hat{\sigma}_z + J\left(\delta\hat{a}_T^\dagger\delta\hat{a}_L + h.c. + \delta\hat{a}_T^\dagger\delta\hat{a}_R + h.c.\right) + \\ & + g(\delta\hat{a}_T\hat{\sigma}_+ + h.c.) + \frac{g_0}{\sqrt{2}}\sqrt{\frac{n_{cav}}{2}}\left(\hat{b}^\dagger + \hat{b}\right)\left[\left(\delta\hat{a}_L^\dagger + \delta\hat{a}_L\right) + \left(\delta\hat{a}_R^\dagger + \delta\hat{a}_R\right)\right] \end{aligned} \quad (S22)$$

where we defined  $\sqrt{n_{cav}} = \sqrt{2}|\bar{a}_R| = \sqrt{2}|\bar{a}_L|$  and we have used the fact that, with the pumping scheme described in the previous section,  $\bar{a}_R = -\bar{a}_L$  and  $\bar{a}_T = 0$ . We note that this Hamiltonian is still equivalent to the original one, since no terms have been neglected. The Hamiltonian has been solved numerically (with the method discussed in sec. I) by either setting  $n_{cav} = 0$  (for the case in which the mode  $\hat{a}_0$  is not externally pumped) or setting a fixed value of  $n_{cav}$  (continuous optical pumping of the system). To address the case of a square-pulse excitation, we first calculate the time evolution of the (classical) cavity fields amplitudes upon the external pumping described above, neglecting the presence of the QE and the mechanical mode. From the amplitudes  $\bar{a}_R(t)$  and  $\bar{a}_L(t)$ , we calculate the population of the mode  $\hat{a}_0$  ( $n_{cav}(t)$ ), which is used as a time-dependent parameter in solving the Hamiltonian in eq. S22.

### IX. INFLUENCE OF CAVITY POPULATION AND DEPHASING ON THE FIDELITY OF THE ONE-PHONON STATE CREATION

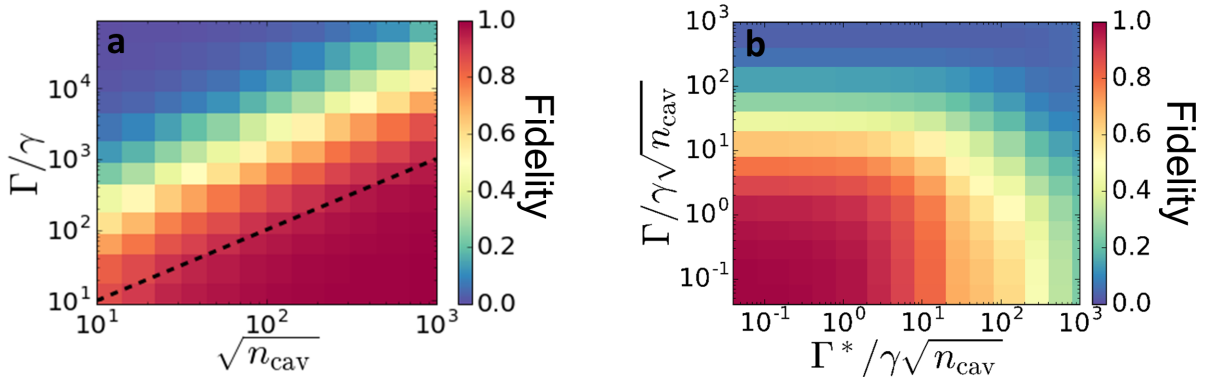


FIG. S5. (a) Fidelity of the creation of the one-phonon state versus the emitter decay rate  $\Gamma$  and the population of the mode  $\hat{a}_0$  ( $n_{cav}$ ). The dashed black line indicates the condition  $\Gamma/\gamma = \sqrt{n_{cav}}$ . (b) Fidelity of the creation of the one-phonon state versus the emitter decay rate  $\Gamma$  and the emitter pure dephasing rate  $\Gamma^*$ .

As mentioned in the main text, the mode field coupling rate can be enhanced by increasing the population of the mode  $\hat{a}_0$ . This is beneficial for, *e.g.*, the fidelity of the emitter-phonon

swapping. The colormap in fig. S5a shows the fidelity of the creation of the one-phonon state (calculated with the transformed Hamiltonian of the three-cavity system, eq. S22) versus the emitter decay rate  $\Gamma$  and the population of the mode  $\hat{a}_0$ . All the other parameters are the same as in the main text. As expected, near-unity fidelity is obtained when  $\gamma\sqrt{n_{\text{cav}}} > \Gamma$ .

Due to the coherent nature of the MFC, pure dephasing of the QE is expected to further decrease the fidelity. As shown in fig. S5b, the QE decay  $\Gamma$  and pure dephasing rates  $\Gamma^*$  play a similar role in determining the fidelity, and  $\Gamma^* \ll \gamma\sqrt{n_{\text{cav}}}$  is required as well to achieve large fidelity. We note that for the system considered for the calculations shown in this work (NV centers in diamond), the pure dephasing rate is typically of the same order of magnitude of the decay rate in bulk [7].

## X. CREATION OF OTHER NONCLASSICAL MECHANICAL STATES

The MFC creates an interaction between a QE and a mechanical resonator whose coupling rate is controllable in time by the optical intensity. This unique characteristic allows us to create, in principle, any nonclassical state of the mechanical resonator, and the fidelity is only limited by the QE and optical losses. In general, as demonstrated by Law and Eberly [8], in any Jaynes-Cumming-like system in which the interaction between a quantum emitter and a bosonic field (*i.e.*, the mechanics in this case) can be controlled in time, arbitrary quantum states of the bosonic field can be created. This is achieved by a sequence of pulses in which the QE is first partially excited from the outside, and then the excitation is transferred to the bosononic field by switching on the coupling [8].

In practice, simple nonclassical states such as larger Fock states or a superposition of Fock states can be easily realized by a sequence of few  $\pi$ - and/or  $\pi/2$ -pulses applied to the QE and the cavity, as shown in fig. 3c of the main text. Here, we first explain the creation of these states within a simplified model in which the MFC is added *ad-hoc* to a system composed of one cavity, one mechanical resonator and a QE (eqs. 1-2 of the main text). Then, we show how the same dynamics, apart for the additional losses due to the other two optical modes, is obtainable with the three-cavity system. We used the same parameters as in the main text,  $\{\omega_c, \Omega_M, g, g_0, J\} = 2\pi \cdot \{4.7 \cdot 10^5, 14, 20, 0.004, 360\}$  GHz, and we consider the case of red-detuned cavity,  $\omega_A = \omega_c + \Omega_M$ . The cavity losses are set to  $\kappa/(2\pi) = 3$  GHz. For simplicity, we neglect mechanical losses. The maximum photon population in the cavity mode upon external pumping is  $\bar{n}_{\text{cav}} = 5 \cdot 10^4$ .

### X.1. One-cavity model with MFC

We consider the simple Hamiltonian of eq. 2 (in which we already performed the linearisation of the optical field). Moreover, we add an external pumping to the atom, we consider only the dominant terms of the tripartite interaction (see discussion in the main text after eq. 1) and we switch to a reference frame rotating at frequency  $\Omega_M = \omega_A - \omega_c$ . The Hamiltonian reads

$$\hat{H} = \gamma\sqrt{n_{\text{cav}}(t)}(\hat{b}\hat{\sigma}_+ + \hat{b}^\dagger\hat{\sigma}_-) + E_A(t)(\hat{\sigma}_+ + \hat{\sigma}_-), \quad (\text{S23})$$

where  $E_A(t)$  is a time-dependent external field resonant with the QE and  $n_{\text{cav}}(t)$  is the population of the optical mode. Similar to the calculations shown before,  $n_{\text{cav}}(t)$  is calculated as the response of the cavity mode to a sequence of square pulses. The value of  $\gamma$  is assumed identical to the one obtainable in the three-cavity system with the same set of parameters, *i.e.*,  $\gamma = gg_0/(2J) = (2\pi) \cdot 0.11$  MHz.

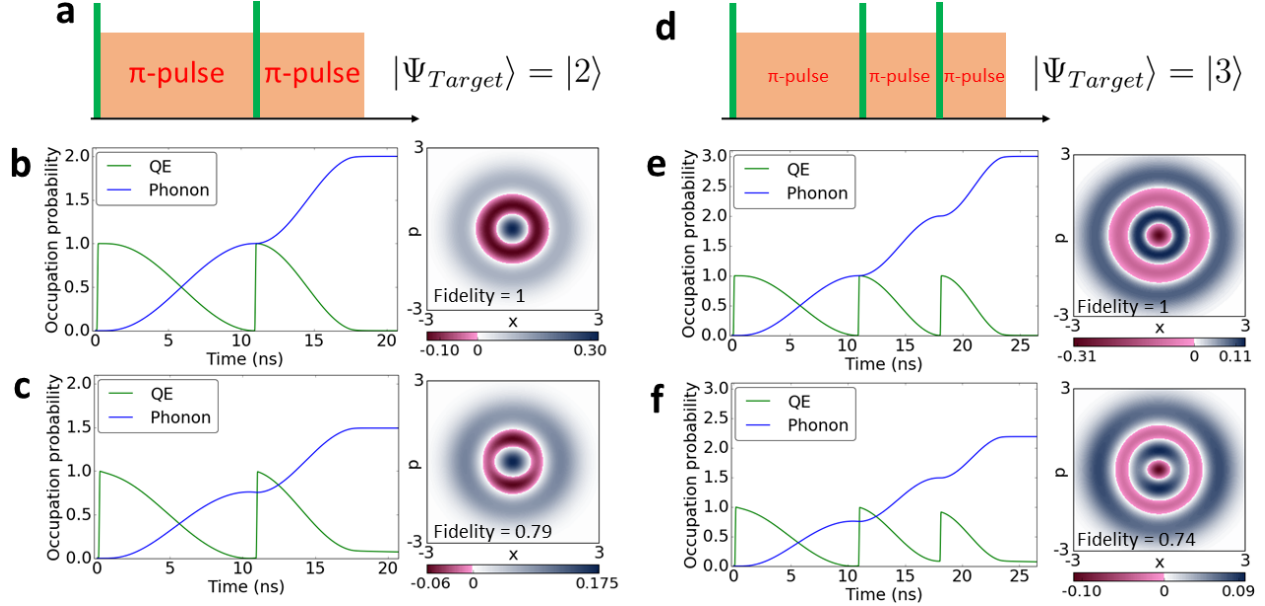


FIG. S6. Creation of larger Fock states with the one-cavity MFC model. (a-c) Creation of the Fock state  $n = 2$ . (a) Protocol used. Green rectangles indicate  $\pi$ -pulses applied to the QE, while red rectangles are pulses applied to the cavity. (b) Time evolution in absence of QE losses (left) and Wigner map of the mechanical state created (right). (c) Same as in (b), but with finite QE losses  $\Gamma/(2\pi) = 0.05$  GHz. (d-f) Same as in (a-c), for the Fock state  $n = 3$ .

Arbitrary mechanical Fock states can be obtained by first exciting the QE and then transferring the excitation to the mechanical resonator by exciting the cavity mode with a  $\pi$ -pulse, *i.e.*, a pulse with a duration  $\Delta t$  and a maximum photon population  $\bar{n}_{\text{cav}}$  such that  $(2\gamma\sqrt{m})\sqrt{\bar{n}_{\text{cav}}}\Delta t = \pi$ , where  $m$  is the phonon number to be achieved at the end of the pulse. In fig. S6 we show the protocols to create the Fock states  $n = 2$  (fig. S6a) and  $n = 3$  (fig. S6d). We calculated the system dynamics upon these pumping protocols for both the lossless case (figs. S6b and S6e) and for the case in which the QE has finite losses identical to the case considered in the main text,  $\Gamma/(2\pi) = 0.05$  GHz (figs. S6c and S6f). For each case we show the temporal evolution of the system (in terms of QE and phonon occupation probability) and the Wigner map of the created mechanical states. The fidelity of the created states are displayed inside the corresponding Wigner maps. For zero emitter losses, the desired nonclassical states are created with unity fidelity, while the fidelity is somewhat reduced when QE losses are introduced. Nonetheless, it is high enough to display the characteristic negativity (nonclassicality) of the Wigner function.

Other nonclassical states, such as the superposition of two or more Fock states, can be obtained by using different pulse sequences. For example, if in the case of the Fock state  $n = 2$  (fig. S6a) the first  $\pi$ -pulse on the cavity is replaced by a  $\pi/2$  pulse, the state  $(|g, 0\rangle - |g, 2\rangle)/\sqrt{2}$  is obtained (fig. S7a), where  $g$  refers to the ground state of the QE and the numbers are the Fock states of the mechanics. The temporal evolution of the system and the Wigner map of the mechanical states are shown in figs. S7b and S7c, for the lossless case and the case with finite QE losses, respectively. Again, the only limitation to the fidelity of the created state is due to the QE losses.

## X.2. Three-cavity model

We now verify that these nonclassical states can be obtained also in the three-cavity system with a large optical pumping. We numerically solve the full, non-linearized, Hamiltonian in eq. S22 (with an additional term to externally pump the QE), in which  $n_{\text{cav}}$  is a time-dependent parameter which describes the optical pumping of the mode  $\hat{a}_0$ . In figs.

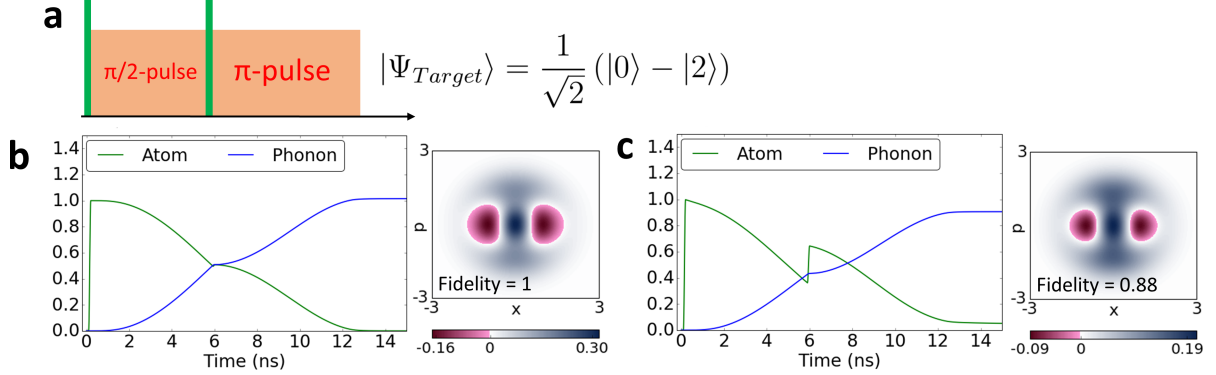


FIG. S7. Creation of the superposition state  $|\Psi_{\text{Target}}\rangle = (|g, 0\rangle - |g, 2\rangle)/\sqrt{2}$  with the one-cavity MFC model. (a) Protocol used. Green rectangles indicate  $\pi$ -pulses applied to the QE, while red rectangles are pulses applied to the cavity. (b) Time evolution in absence of QE losses (left) and Wigner map of the mechanical state created (right). (c) Same as in (b), but with QE losses  $\Gamma/(2\pi) = 0.05$  GHz.

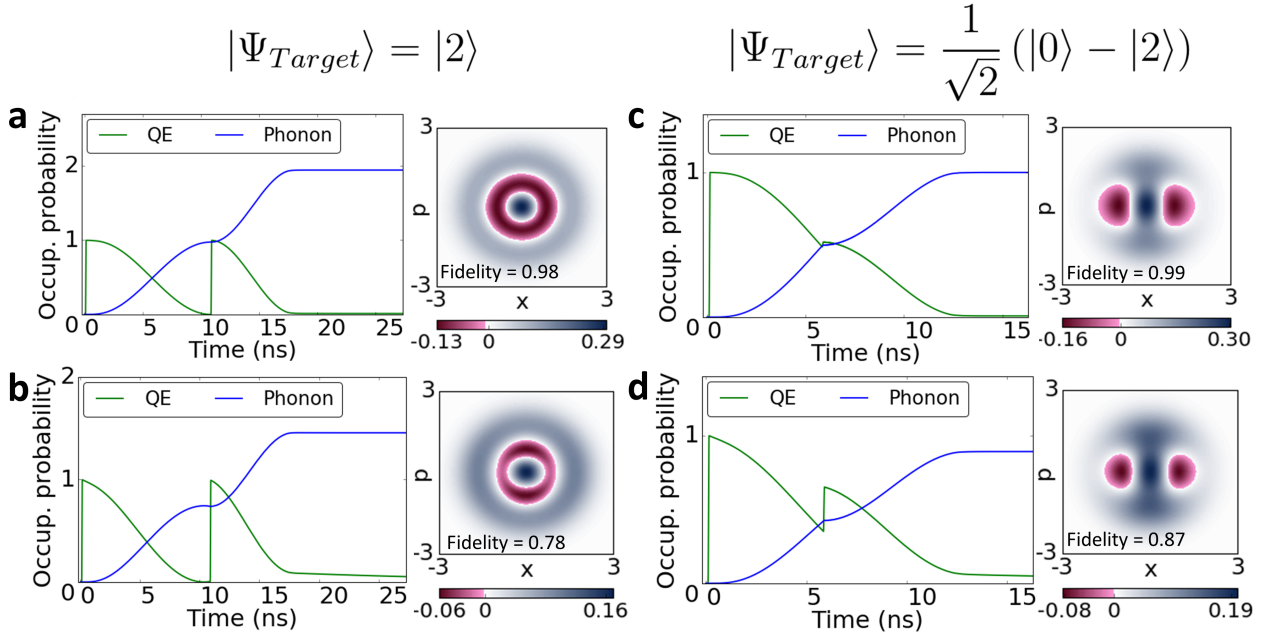


FIG. S8. Creation of nonclassical mechanical states with the three-cavity system. (a-b) Creation of the Fock state  $n = 2$  (same pumping protocols as fig. S6a). (a) Time evolution in absence of QE losses (left) and Wigner map of the mechanical state created (right). (b) Same as in (a), but with QE losses  $\Gamma/(2\pi) = 0.05$  GHz. (c-d) Creation of the superposition state  $|\Psi_{\text{Target}}\rangle = (|g, 0\rangle - |g, 2\rangle)/\sqrt{2}$  (same pumping protocols as fig. S7a). Panels (c) and (d) show the same plots as in (a-b), but for the superposition state.

S8(a-b) we show the creation of the Fock state  $n = 2$ , without (panel a) and with (panel b) emitter losses ( $\Gamma/(2\pi) = 0.05$  GHz). The pumping protocol used is the same as in fig. S6a. The only difference with respect to the one-cavity system is that, even for zero QE losses, the fidelity (0.98) is lower than one, as a result of the coupling of the QE with the other two optical supermodes. Similar results are obtained in the creation of the superposition state  $|\Psi_{\text{Target}}\rangle = (|g, 0\rangle - |g, 2\rangle)/\sqrt{2}$  (figs. S8(c-d)) that, in absence of QE losses (fig. S8c), is created with fidelity 0.99. While reducing the cavity losses ( $\kappa$ ) will reduce the QE decay induced by the other two optical supermodes ( $\Gamma^{(\pm)}$ , see sec. VI), it will also increase the temporal duration of the cavity response to the external pumping. This results in a longer duration of the excitation transfer between the QE and the phonon, and therefore the QE decay ( $\Gamma$ ) will have a larger influence on the fidelity. This trade-off leads to an optimal value of  $\kappa$  that maximizes the fidelity, which is numerically found to be  $\kappa_{\text{opt}}/(2\pi) \approx 3$  GHz in this case.

## XI. ROLE OF THE OTHER TWO MECHANICAL SUPERMODES

In the main text, and in the previous sections of this supplemental information, we assumed that only the mechanical supermode of interest is relevant for the system dynamics. To verify this assumption, we performed additional numerical calculations based on the Hamiltonian in eq. S10, *i.e.* by considering explicitly the three mechanical uncoupled modes.

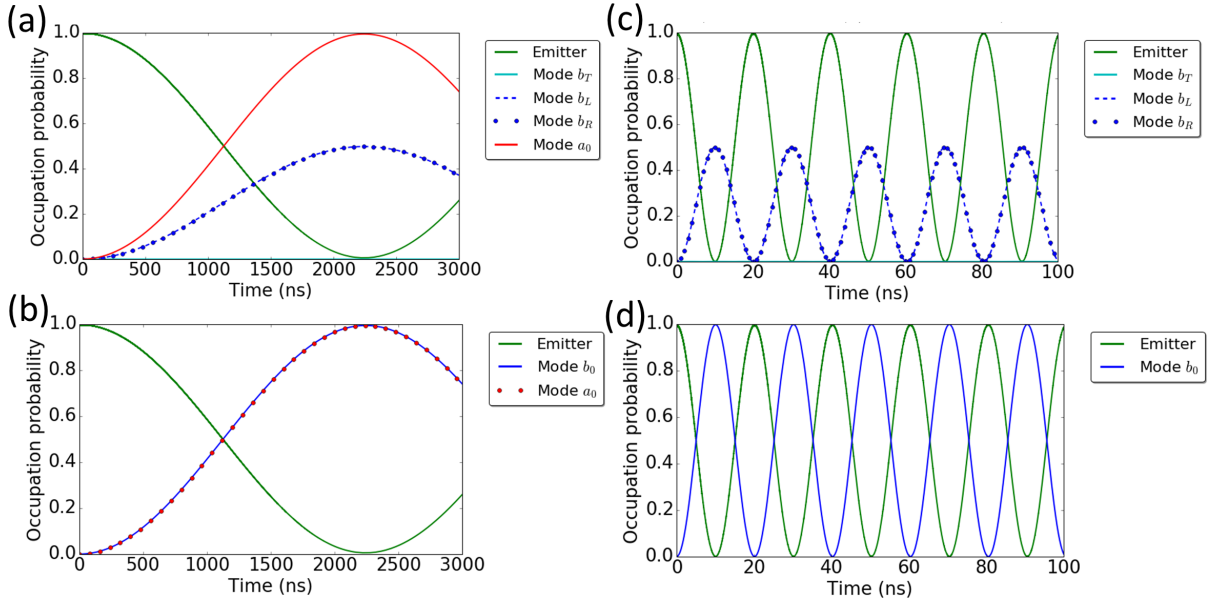


FIG. S9. Numerical calculation of the full Hamiltonian in eq. S10, *i.e.* considering all mechanical modes. The parameters of the system are the same as in fig. 3a of the main text. (a-b) Vacuum oscillations of the system (*i.e.* the optical cavities are not pumped). (a) The mechanical modes  $\hat{b}_L$  and  $\hat{b}_R$  are equally populated (blue solid and blue dotted lines) and reach a maximum occupation probability of 0.5, while the mechanical mode  $\hat{b}_T$  is not populated (cyan solid line). (b) When switching to the supermode basis (through eqs. S11), the mode  $\hat{b}_0$  behaves exactly as shown in previous calculations in which the other supermodes were neglected. (c-d) Same as in panels (a-b) but with an external pumping of the optical cavities, such that the optical supermode  $\hat{a}_0$  contains  $n = 5 \cdot 10^4$  photons.



We use the same parameters as in the main text, and a mechanical coupling  $J_M/(2\pi) = 50$  MHz. Initially, we assume no mechanical losses ( $\Gamma_M = 0$ ). After calculating the time evolution of the modes  $\hat{b}_L$ ,  $\hat{b}_R$  and  $\hat{b}_T$ , we derive the time evolution of the three mechanical supermodes based on eqs. S11. As shown in fig. S9, both in the unpumped and the pumped case the dynamics of the mechanical mode  $\hat{b}_0$  matches that found previously, where the other mechanical modes were neglected. The population of the other two mechanical supermodes ( $\hat{b}_+$  and  $\hat{b}_-$ ) is always zero (not shown in these plots). This is due to the fact that these other mechanical supermodes interact with the optical cavities in a way which does not lead to any tripartite interaction (see eq. S12), and their presence is therefore negligible for proper choice of frequencies.

Interestingly, these results hold true even for large mechanical losses  $\Gamma_M \gg J_M$ , and the fidelity of the one-phonon-state creation does not depend on the mechanical interaction  $J_M$ , as we show in fig. S10a. Decrease of the fidelity is only observed if the mechanical losses become larger than  $\gamma\sqrt{n_{\text{cav}}}$ . The fact that the the presence of the other two mechanical supermodes can be completely neglected is also confirmed by fig. S10b: here, we compare the phonon population of the mode  $\hat{b}_0$  versus  $\Gamma_M$  for the case of a three-cavity system (*i.e.* a horizontal cut of panel a) with the same graph calculated for a system composed by one optical cavity and one mechanical resonator, where the MFC is introduced *ad-hoc* (eq. 2 of the main text). The identical behaviours indicate that a large  $\Gamma_M$  introduces only additional losses but does not make the system interact with the other mechanical supermodes. The small deviation between the two curves for  $\Gamma_M/(\gamma\sqrt{n_{\text{cav}}}) < 1$  is due to the additional losses present in the three-cavity system due to the optical modes  $\hat{a}_{\pm}$ .

From a fundamental point of view, therefore, there are no constraints on the mechanical coupling  $J_M$ . Nonetheless, for practical reasons it could be beneficial to work in a regime where  $J_M > \Gamma_M$ , such that the mode of interest is spectrally separated from the other mechanical supermodes and it can be individually addressed in other ways for, *e.g.*, cooling or readout.

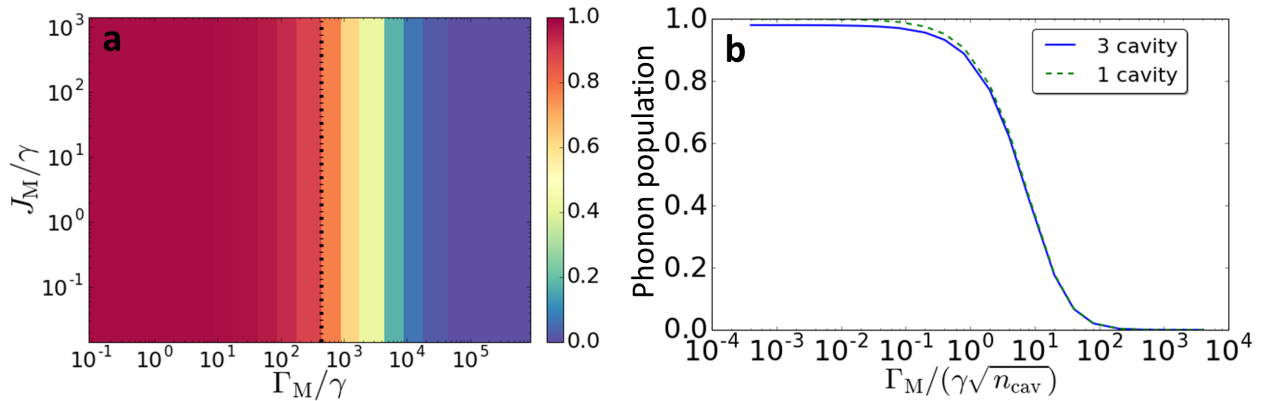


FIG. S10. Calculations of three-cavity system without neglecting the other mechanical supermodes. The optical mode  $\hat{a}_0$  is continuously pumped with  $n_{\text{cav}} = 5 \cdot 10^4$  and the QE is initially excited. All the other parameters are the same as in the main text and the QE losses have been neglected for simplicity. (a) Maximum phonon population (of the mode  $\hat{b}_0$ ) achieved during the system evolution, versus the mechanical interaction  $J_M$  and the mechanical losses  $\Gamma_M$ . The dashed-dotted vertical line indicates the condition  $\Gamma_M = \gamma\sqrt{n_{\text{cav}}}$ . (b) Horizontal cut of panel (a) for the lowest value of  $J_M$  (solid blue line) compared with the same quantity calculated in a MFC model with one cavity and one resonator (green dashed line, see eq. 2 of main text).



## XII. DESIGN OF THE PROPOSED STRUCTURE

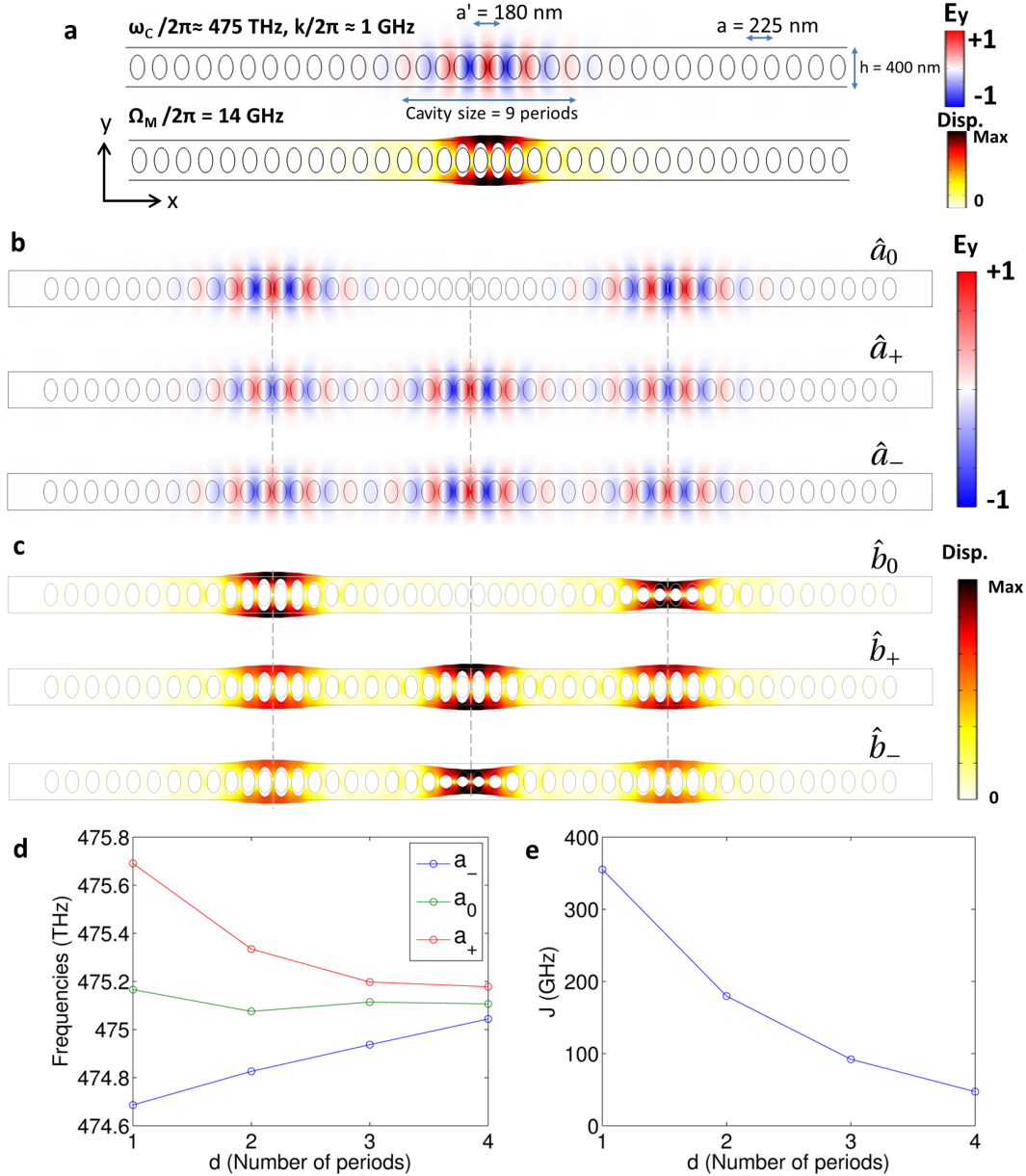


FIG. S11. Numerical investigation of the proposed structure (see text for details). (a) Electric field pattern (top) and mechanical displacement pattern (bottom) of the colocalized optical and mechanical mode for a single cavity. The  $y$ -component of the electric field is shown. The mode frequencies and the optical losses are indicated in the figure. (b-c) Three identical cavities on a nanobeam, separated by 2 periods. The position of the three cavities is marked by the vertical grey dashed lines. (b) Electric field pattern of the three optical supermodes when the mechanical mode is at rest. (c) Displacement pattern of the three mechanical supermodes. (d) Frequencies of the three optical supermodes as a function of the cavities separation (*i.e.* number of periods). (e) Values of the optical interaction  $J$  (extrapolated from the plot in panel (d), see text) versus the cavity separation.

The structure proposed in the main text is composed of three in-line defect cavities in a

photonic crystal diamond nanobeam. Simulations have been performed with a finite element method commercial software (COMSOL). We first designed the single cavity structure (fig. S11a), which is realized by quadratically tapering the lattice constant across 9 lattice periods, from the unperturbed value of  $a = 225$  nm to the central value of  $a' = 0.8a = 180$  nm. The design is inspired by the one proposed recently by Lee *et al.*[4]. The holes have elliptical shape, with semiaxes equal to  $a_x = 0.32a$  and  $a_y = 0.55a$ . The width of the nanobeam (along the  $y$ -direction) is  $w = 400$  nm, and the thickness is 200 nm. This structure supports an optical mode with frequency  $\omega_c/2\pi = 475$  THz and  $Q \approx 3 \cdot 10^5$  (fig. S11a, top panel), and a co-localized mechanical mode with frequency  $\Omega_M/2\pi = 14$  GHz (fig. S11a, bottom panel). The dispersive optomechanical coupling between the optical and mechanical modes has been calculated by evaluating the moving-boundary and photoelastic contribution separately (as discussed in ref. 9) and amounts to  $g_0/2\pi \approx 4$  MHz. We notice that we did not perform any systematic optimization on our design, and therefore the optical Q factor and optomechanical coupling rate could be further increased by carefully optimizing the design parameters (*e.g.*, by tapering also the holes' semiaxes) as discussed by other authors [9].

After having chosen a design for the single cavity, we considered three identical defect cavities on the same nanobeam, separated by an equal distance  $d$ , defined as the number of unperturbed periods between two adjacent cavities. For each distance  $d$ , we calculated the optical and mechanical supermodes. Figure S11b shows the  $E_y$  field component for the three optical supermodes. Note that the optical mode of interest,  $\hat{a}_0$ , has no field in the central cavity, differently from the other two modes. The modes  $\hat{a}_+$  and  $\hat{a}_-$  differ in the relative sign between the field in the lateral cavities and the field in the central cavity, as expected from the results of the coupled-mode theory (eq. S14). Figure S11c shows the displacement pattern of three mechanical supermodes. Similarly to the optical case, the mechanical supermodes  $\hat{b}_+$  and  $\hat{b}_-$  have opposing oscillation phase in the central and the lateral cavities. Figure S11d shows the frequencies of the three optical supermodes versus the cavity separation  $d$ . The frequency of the mode  $\hat{a}_0$  is expected to be independent of the cavity interaction  $J$  (and therefore the cavity distances). The small frequency deviations observed can be due to either the finite mesh size (which makes the cavities slightly different from each other) or to the breakdown of the coupled-mode theory for very short cavity distances. We notice that this does not affect our theoretical model, since the mechanical movement does not change the rate  $J$ . In fig. S11e we show the estimated optical interaction rate  $J$  as a function of the inter-cavity distance. For each distance  $d$ ,  $J$  has been estimated by the formula  $|\omega_+ - \omega_-| = 2\sqrt{2}J$ .

In fig. 2I of the main text we show the expected pattern of the mode  $\hat{a}_0$  for the case in which the system is mechanically perturbed, such that the lateral cavities are detuned by  $\Delta/J = 0.5$ . The shown electric field pattern has been calculated analytically from the formula

$$E_{0,\Delta}(x, y) = C_0(\Delta)E_0(x, y) + C_+(\Delta)E_+(x, y) + C_-(\Delta)E_-(x, y), \quad (\text{S24})$$

where  $E_i(x, y)$ , with  $i = \{0, +, -\}$ , are the electric field patterns of the supermodes when the mechanical mode is at rest (fig. S11b), and the coefficients  $C_i(\Delta)$  are calculated analytically from the coupled mode theory.

### XIII. TRIPARTITE INTERACTION IN A FABRY-PEROT CAVITY

The interaction described in this work relies on a variation of the electric field at the emitter position upon a mechanical displacement. A small field variation occurs in principle also in a simple Fabry-Perot (FP) cavity when one of the mirrors is displaced by a mechanical

resonator [10–12]. However, the interaction strength in this case is much smaller than the one obtained in the two and three cavity systems. In particular, for an emitter placed in a node of the  $n$ -th FP mode, the tripartite coupling rate reads  $\gamma_n = \pi g g_0 / \omega_1$ , where  $g$  is the maximum Rabi coupling between the QE and the  $n$ -th mode,  $g_0$  is the dispersive coupling induced by the mirror movement and  $\omega_1$  is the fundamental cavity frequency. The coupling rate bears similarity with the one derived for the two- and three-cavity systems, but with the important difference that the role of the intercavity interaction rate  $J$  is taken by  $\omega_1$ , which, in the FP cavity, corresponds also to the frequency spacing between the unperturbed modes. This, in the visible and near-IR regime, severely limits the achievable values of  $\gamma_n$ , especially since any effort to reduce  $\omega_1$  reduces both  $g$  and  $g_0$ .

- 
- [1] J. Johansson, P. Nation, and F. Nori, *Comput. Phys. Commun.* **183**, 1760 (2012).
  - [2] M. A. Nielsen and I. L. Chuang, Cambridge: Cambridge University Press **2**, 23 (2000).
  - [3] M. Ludwig, A. H. Safavi-Naeini, O. Painter, and F. Marquardt, *Phys. Rev. Lett.* **109**, 063601 (2012).
  - [4] J. C. Lee, D. O. Bracher, S. Cui, K. Ohno, C. A. McLellan, X. Zhang, P. Andrich, B. Alemán, K. J. Russell, A. P. Magyar, I. Aharonovich, A. B. Jayich, D. Awschalom, and E. L. Hu, *Appl. Phys. Lett.* **105**, 261101 (2014).
  - [5] A. Faraon, C. Santori, Z. Huang, V. M. Acosta, and R. G. Beausoleil, *Phys. Rev. Lett.* **109**, 033604 (2012).
  - [6] P. Lodahl, S. Mahmoodian, and S. Stobbe, *Rev. Mod. Phys.* **87**, 347 (2015).
  - [7] L. Robledo, H. Bernien, I. van Weperen, and R. Hanson, *Phys. Rev. Lett.* **105**, 177403 (2010).
  - [8] C. Law and J. Eberly, *Phys. Rev. Lett.* **76**, 1055 (1996).
  - [9] M. J. Burek, J. D. Cohen, S. M. Meenehan, T. Ruelle, S. Meesala, J. Rochman, H. A. Atikian, M. Markham, D. J. Twitchen, M. D. Lukin, O. Painter, and M. Lončar, *Optica* **3**, 1404 (2016).
  - [10] S. Barzanjeh, M. H. Naderi, and M. Soltanolkotabi, *Phys. Rev. A* **84**, 063850 (2011).
  - [11] Y. Chang, H. Ian, and C. Sun, *J. Phys. B* **42**, 215502 (2009).
  - [12] W. Wang, L. Wang, and H. Sun, *J. Korean Phys. Soc.* **57**, 704 (2010).