

# Molecular Platform for Frequency Upconversion at the Single-Photon Level

Philippe Roelli,<sup>1</sup> Diego Martin-Cano<sup>2</sup>,<sup>✉</sup> Tobias J. Kippenberg,<sup>1,\*</sup> and Christophe Galland<sup>1,†</sup>

<sup>1</sup>*EPFL, Swiss Federal Institute of Technology, Institute of Physics, Lausanne, Switzerland*

<sup>2</sup>*Max Planck Institute for the Science of Light, Erlangen, Germany*



(Received 21 October 2019; revised 23 March 2020; accepted 23 June 2020; published 14 September 2020)

Direct detection of single photons at wavelengths beyond  $2\ \mu\text{m}$  under ambient conditions remains an outstanding technological challenge. One promising approach is frequency upconversion into the visible (VIS) or near-infrared (NIR) domain, where single-photon detectors are readily available. Here, we propose a nanoscale solution based on a molecular optomechanical platform to up-convert photons from the far- and mid-infrared (covering part of the terahertz gap) into the VIS-NIR domain. We perform a detailed analysis of its outgoing noise spectral density and conversion efficiency with a full quantum model. Our platform consists in doubly resonant nanoantennas focusing both the incoming long-wavelength radiation and the short-wavelength pump laser field into the same active region. There, infrared active vibrational modes are resonantly excited and couple through their Raman polarizability to the pump field. This optomechanical interaction is enhanced by the antenna and leads to the coherent transfer of the incoming low-frequency signal onto the anti-Stokes sideband of the pump laser. Our calculations demonstrate that our scheme is realizable with current technology and that optimized platforms can reach single-photon sensitivity in a spectral region where this capability remains unavailable to date.

DOI: [10.1103/PhysRevX.10.031057](https://doi.org/10.1103/PhysRevX.10.031057)

Subject Areas: Condensed Matter Physics,  
Photonics Plasmonics

## I. INTRODUCTION

Many applications in security, material science, and health care would benefit from the development of new technologies for far- and mid-infrared (FIR and MIR, respectively) detection and thermal imaging [1]. Driven by applications in astronomy, novel cryogenic detectors in the FIR range appeared in the past few years [2,3]. However, the ability to efficiently manipulate such electromagnetic signals at room temperature is still lacking [4,5]. In particular, single-photon detection, which is now routine in the visible (VIS) and near-infrared (NIR) region (wavelength in vacuum from 400 to 2000 nm), remains impossible or unpractical at longer wavelengths. The development of new detection devices operating without complex cryogenic apparatus, and featuring improved sensitivity, lower noise, and a reduced footprint, would significantly

impact sensing, imaging, spectroscopy, and communication technologies.

In this work, we propose a new route to achieve low-noise detection of noncoherent radiation between 5 and 50 THz by leveraging optomechanical transduction with molecules [6], whose natural oscillation frequencies are resonant with the incoming field. Our strategy consists in converting the incoming low-frequency signal onto the anti-Stokes sideband of a pump laser in the VIS-NIR domain, where detectors with single-photon sensitivity are readily available [7,8]. This approach is inspired by the recent realization of coherent frequency conversion using different types of optomechanical cavities [9–15] and is conceptually distinct from a recently demonstrated detection scheme assisted by a microfabricated resonator [16]. As an outlook, we propose to leverage constructive interference between signals coming from an array of coherently pumped up-converters in order to increase further the strength of the converted signal over the incoherent thermal noise.

While coherent conversion from the MIR to the VIS-NIR domain has so far been achieved by sum-frequency generation in bulk nonlinear crystals [17–20], these schemes operate under several watts of pump power and require phase matching between the different fields propagating in the crystal. Our scheme, on the contrary, relies solely on the spatial overlap of the two incoming fields. Indeed, we use a nanometer-scale dual antenna that confines both

\*Corresponding author.

tobias.kippenberg@epfl.ch

†Corresponding author.

chris.galland@epfl.ch

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

electromagnetic fields into similar subwavelength mode volumes. The optomechanical interaction with the vibrational system takes place in the near field, without the need to fulfill a phase-matching condition. Moreover, thanks to the giant field enhancement provided by plasmonic nanogaps, the required pump power to achieve efficient conversion is dramatically reduced.

The protocol that we introduce leverages the intrinsic ability of specific molecular vibrations to interact both resonantly with MIR-FIR fields and parametrically with VIS-NIR fields, as routinely observed in infrared absorption and Raman spectroscopy, respectively. The wealth of accessible vibrational modes and frequencies [21,22] offers a convenient toolbox to realize efficient frequency upconversion in the technologically appealing region of thermal imaging.

We first introduce the framework describing the interaction between a molecular vibration and two electromagnetic fields, one that is resonant with the vibrational frequency and the other one that is parametrically coupled to it through the molecular polarization. We compute the conversion efficiency and the noise figures of merit of our novel device as a function of the optical pump detuning and power. We illustrate the achievable performance with a device operating at 30 THz (10  $\mu\text{m}$ ) and find internal conversion efficiencies on the order of a few percent and noise-equivalent power below  $10^{-12}$  W/ $\sqrt{\text{Hz}}$ . Finally, we demonstrate how our approach may be used to reach single-photon detection at frequencies down to approximately 5 THz.

## II. OPTICAL CONVERSION SCHEME

We start with the description of the two types of interactions leveraged in the conversion process and describe the relevant parameters. For simplicity, we now use the abbreviation IR to denote MIR or FIR fields, depending on the vibrational frequency considered. First, we model the resonant absorption process. We assume that the vibrational system is weakly driven, meaning that the average number of excited collective vibrational quanta is much smaller than the total number of molecular oscillators coupled to the incoming field,  $N_{\text{IR}}$ . At the single-molecule level, this easily satisfied condition corresponds to neglecting transitions beyond the ground and first excited vibrational states. The collective excitation of an ensemble of vibrational modes can thus be treated as an ensemble of two-level systems [23].

The interaction part of the Hamiltonian is correspondingly approximated by

$$\hat{H}_{\text{int}} = -i\hbar g_{\text{IR},0}^{(N_{\text{IR}})} \sqrt{\bar{n}_{\text{IR}}} (\hat{a}_{\text{IR}}^\dagger \hat{\sigma}_\nu^- + \hat{a}_{\text{IR}} \hat{\sigma}_\nu^+), \quad (1)$$

with  $\hat{a}_{\text{IR}}^\dagger$  and  $\hat{a}_{\text{IR}}$  the IR field bosonic ladder operators and  $\hat{\sigma}_\nu^+$  and  $\hat{\sigma}_\nu^-$  the raising and lowering operators, respectively, of the collective two-level system described by a transition

frequency  $\omega_\nu$ .  $g_{\text{IR},0}^{(N_{\text{IR}})} = \sqrt{N_{\text{IR}}} g_{\text{IR},0}$  is the collective resonant vacuum coupling rate of the vibrational mode  $\nu$  for  $N_{\text{IR}}$  identical molecules, and  $\bar{n}_{\text{IR}}$  the mean occupation of the IR antenna mode.

The incoming IR field at frequency  $\omega_{\text{IR}}$  is enhanced by a frequency-matched antenna and performs work on the collective transition dipole  $\vec{d}_\nu$  of the molecular vibration [24]. On resonance ( $\omega_{\text{IR}} = \omega_\nu$ ), the average number of created phonons is (see the Appendix C for a detailed derivation)

$$\bar{n}_b^{\text{IR}} = \left( \frac{2g_{\text{IR},0}^{(N_{\text{IR}})}}{\Gamma_{\text{tot}}} \right)^2 \frac{\eta_{\text{IR}}}{\kappa_{\text{IR}}^{\text{IR}}} |\langle \hat{a}_{\text{IR}}^{\text{in}} \rangle|^2 \quad (2)$$

with  $|\langle \hat{a}_{\text{IR}}^{\text{in}} \rangle|^2$  the incoming IR photon flux. In this expression,  $\kappa_{\text{IR}}^{\text{IR}} = \kappa_{\text{ex}}^{\text{IR}} + \kappa_0^{\text{IR}}$  is the loss rate of the antenna at the incoming frequency, which is the sum of the external decay rate  $\kappa_{\text{ex}}^{\text{IR}}$  (by radiative coupling to far-field modes) and the internal decay rate  $\kappa_0^{\text{IR}}$  (by absorption in the metal).  $\eta_{\text{IR}} = \kappa_{\text{ex}}^{\text{IR}} / \kappa_{\text{IR}}^{\text{IR}}$  is the coupling ratio of the antenna and  $\Gamma_{\text{tot}}$  the total vibrational decay rate, where the intrinsic vibrational linewidth  $\Gamma_\nu$  is modified by its coupling to the IR antenna [25] and the optomechanical interaction with the pump laser, as explained below.

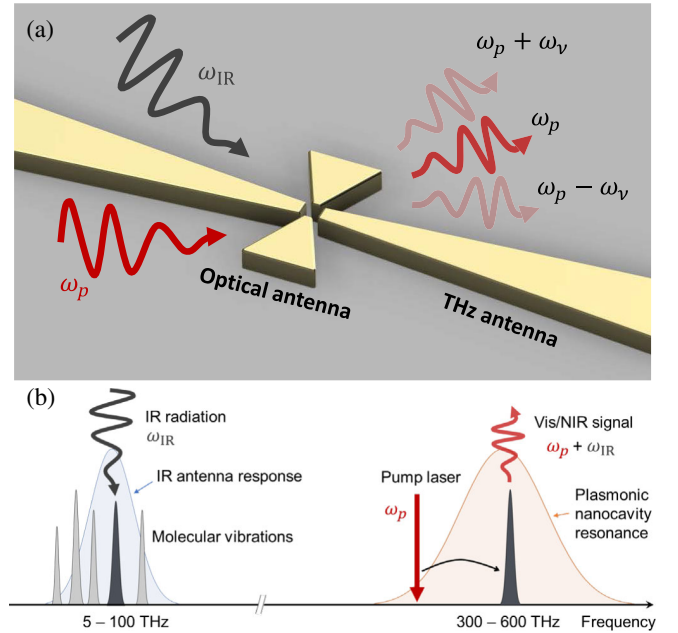


FIG. 1. (a) Illustration of the envisioned upconversion device. Both electromagnetic modes are collected with the help of the dual antenna and confined within a volume where molecules are located. (b) Frequency picture of the optomechanical conversion mechanism involving both IR absorption and Raman scattering by specific vibrational modes. Here, the pump tone ( $\omega_p$ ) is red-detuned from the optical resonance ( $\omega_c \simeq \omega_p + \omega_{\text{IR}}$ ), while the incoming IR signal is resonant with a specific vibrational mode ( $\omega_{\text{IR}} = \omega_\nu$ ).

As pictured in Fig. 1, we employ a second antenna resonant at  $\omega_c$  (a frequency in the VIS-NIR domain, which we call “optical” domain from here on for brevity), whose decay rates  $\kappa_{\text{ex}}^{\text{opt}}$  and  $\kappa_0^{\text{opt}}$  are defined in the same way as the IR antenna parameters. The optical antenna enhances the parametric optomechanical interaction of the molecular vibration with a pump laser in the optical domain, as described in Ref. [6]. Concisely, the interaction between an optical field and  $N_{\text{opt}}$  molecular oscillators leads to a dispersive interaction described by the Hamiltonian  $\hat{H}_{\text{int}} = -\hbar g_{\text{opt},0}^{(N_{\text{opt}})} \hat{a}_{\text{opt}}^\dagger \hat{a}_{\text{opt}} (\hat{b}_\nu + \hat{b}_\nu^\dagger)$  with  $g_{\text{opt},0}^{(N_{\text{opt}})} = \sqrt{N_{\text{opt}}} g_{\text{opt},0}$  the collective optomechanical vacuum coupling rate and  $\hat{a}_{\text{opt}}^\dagger$  and  $\hat{a}_{\text{opt}}$  ( $\hat{b}_\nu^\dagger$  and  $\hat{b}_\nu$ ) the optical pump field bosonic ladder operators (the vibrational phononic operators at frequency  $\nu$ ). The number of molecules  $N_{\text{opt}}$  participating in the optomechanical interaction with the pump laser may be different from the number  $N_{\text{IR}}$  participating in IR absorption, as elaborated in Sec. III below (cf. also Appendix D5 and D6).

The optical antenna field can be split into an average coherent amplitude  $\alpha$  and a fluctuating term so that  $\hat{a}_{\text{opt}} = \alpha + \delta\hat{a}_{\text{opt}}$ . Expanding to first order in  $\alpha$  the optomechanical interaction, we obtain the linearized interaction

$$\hat{H}_{\text{lin}} = -\hbar g_{\text{opt},0}^{(N_{\text{opt}})} \sqrt{\bar{n}_{\text{opt}}} (\delta\hat{a}_{\text{opt}}^\dagger + \delta\hat{a}_{\text{opt}}) (\hat{b}_\nu + \hat{b}_\nu^\dagger), \quad (3)$$

with  $\bar{n}_{\text{opt}} = |\alpha|^2$  the mean occupation of the optical antenna mode (see Appendix A).

The spectral density of the output field on the optical port in photons/(Hz · s) can be evaluated through the calculation of the two-time correlations of the optical output field operators [26,27]:

$$S^{\text{out}}(\omega) \propto \frac{\kappa_{\text{ex}}^{\text{opt}}}{2\pi} \int_{-\infty}^{\infty} d\tau e^{i\omega\tau} \langle \delta\hat{a}_{\text{opt}}^\dagger(\tau) \delta\hat{a}_{\text{opt}}(0) \rangle. \quad (4)$$

Following previous works in optomechanics [28] and their extension to molecular optomechanics [29,30], we can write an analytical expression of the outgoing spectral density at the anti-Stokes sideband  $S^{\text{out}}(\omega_{\text{aS}}) \propto A^- \bar{n}_f / (\Gamma_\nu^* + \Gamma_{\text{opt}})$  [at the Stokes sideband  $S^{\text{out}}(\omega_{\text{S}}) \propto A^+ (\bar{n}_f + 1) / (\Gamma_\nu^* + \Gamma_{\text{opt}})$ ] with  $\bar{n}_f$  the mean final phonon number of the vibrational mode. Because of the IR and optomechanical interactions, the intrinsic vibrational damping rate is modified:  $\Gamma_{\text{tot}} = \Gamma_\nu^* + \Gamma_{\text{opt}}$  with  $\Gamma_\nu^*$  the IR antenna-assisted damping rate and  $\Gamma_{\text{opt}} = A^- - A^+$  the additional damping rate of electromagnetic origin characterized by the imbalance between the optical antenna-assisted transition rates to the ground  $A^-$  and excited vibrational states  $A^+$  (see Appendix A).

The final phonon number in the vibrational mode,  $\bar{n}_f$ , is given by the expression [26]

$$\bar{n}_f = \frac{\Gamma_\nu^*}{\Gamma_\nu^* + \Gamma_{\text{opt}}} \bar{n}_b + \frac{A^+}{\Gamma_\nu^* + \Gamma_{\text{opt}}}, \quad (5)$$

where  $\bar{n}_b = \bar{n}_b^{\text{IR}} + \bar{n}_{\text{th}}$  is the total phonon number in the absence of optical drive. It is the incoherent sum of the IR-induced vibrational excitation [Eq. (2)] and the thermal noise,  $\bar{n}_{\text{th}} = 1/(\exp[\hbar\omega_\nu/k_B T_{\text{bath}}] - 1)$  for a bath temperature  $T_{\text{bath}}$ . We assume here that the pump laser does not lead to significant Ohmic heating of the system. It is, however, straightforward to model laser-induced heating by introducing a pump-power-dependent bath temperature  $T_{\text{bath}}$ .

The resulting spectral density  $S_{\text{opt}}^{\text{out}}$  in the absence of incoming IR radiation ( $\bar{n}_b^{\text{IR}} = 0$ ) should be integrated over the device’s operational bandwidth ( $\text{BW} \equiv \Gamma_{\text{tot}}$ ) to obtain its dark-count rate  $\tilde{S}_{\text{opt}}^{\text{out}} = \int_{\text{BW}} S_{\text{opt}}^{\text{out}} d\omega$ . The dark-count rate arising from the thermal contribution to the first term in Eq. (5) can be reduced by cooling the bath, whereas the second term describes a minimal noise level resulting from phonon creation by spontaneous Stokes scattering of the pump laser, a process equivalent to quantum backaction in cavity optomechanics. Therefore, an optimal power that maximizes the signal-to-noise ratio (SNR) exists, akin to the standard quantum limit in position measurements.

From these expressions, we are also able to describe the conversion efficiency from an incoming rate of IR photons coupled to the antenna into an outgoing rate of optical photons emitted by the antenna into free space, as  $\tilde{S}_{\text{IR} \rightarrow \text{opt}}^{\text{out}} = \eta_{\text{ext}} |\langle \hat{a}_{\text{IR}}^{\text{in}} \rangle|^2$ , where  $\eta_{\text{ext}} = \eta_{\text{opt}} \cdot \eta_{\text{int}} \cdot \eta_{\text{IR}}$  is an external conversion efficiency [31] with  $\eta_{\text{opt}} = \kappa_{\text{ex}}^{\text{opt}} / \kappa^{\text{opt}}$ .

The internal conversion efficiency  $\eta_{\text{int}}$  can, in turn, be divided into a power-dependent part  $\eta_{\text{OM}}(\bar{n}_{\text{opt}})$  and a part describing the spatial overlap between the IR near field, the optical near field, and the molecular ensemble, which we write as  $\eta_{\text{overlap}}$ . To approximate this last term, we factorize it into two contributions: the spatial overlap between the IR and optical near fields ( $\eta_{\text{mode}}$ ) and the vectorial overlap between the near-field polarization (typically normal to the antenna surface) and the molecular orientation, which we name  $\eta_{\text{pol}}$ ; so that we can write

$$\eta_{\text{int}} \simeq \eta_{\text{pol}} \cdot \eta_{\text{mode}} \cdot \eta_{\text{OM}}(\bar{n}_{\text{opt}}). \quad (6)$$

The evaluation of  $\eta_{\text{overlap}} = \eta_{\text{pol}} \eta_{\text{mode}}$  is detailed in Appendix D5 and D6. The power and detuning dependence of the optomechanical efficiency term  $\eta_{\text{OM}}$  are depicted in Fig. 3(b), and its exact calculation is given in Appendix A.

### III. MOLECULAR TRANSDUCER

The electric dipole moment  $\vec{\mu}_\nu$  and polarizability  $\alpha_\nu$  of a vibrational mode can be extracted from experimental data of resonant light absorption and inelastic light scattering, respectively [21,22]. For vibrational modes lacking centrosymmetry, the derivatives of both quantities with respect



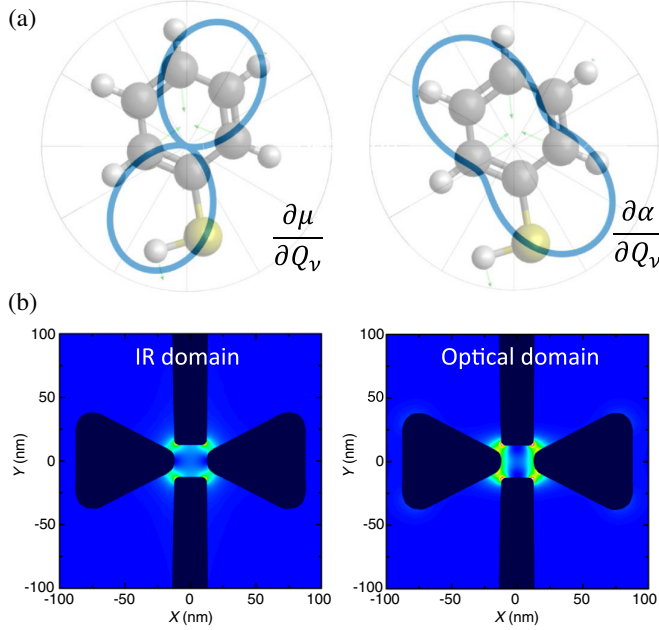


FIG. 2. (a) Polar plots of both the electric moment derivative (left) and the projection of the Raman tensor on the main plane of the molecule (right) for the vibrational mode  $\tilde{\nu} = 1002 \text{ cm}^{-1}$  of the thiophenol molecule (background image). (b) Local density of states distribution inside the dual antenna for an IR mode at 32 THz (left) and a NIR mode at 374 THz (right).

to the displacement coordinate can be nonvanishing [32]. We show such a situation in Fig. 2, where we plot the projections of the derivatives of the electric moment and of the polarizability with respect to the molecular coordinate  $Q_\nu$  of the  $1002 \text{ cm}^{-1}$  mode of thiophenol, which we choose as an example in our calculations. We note that the projection of the tensor  $(\partial\alpha_\nu/\partial Q_\nu)$  onto an axis perpendicular to the principal axis of the electronic moment derivative  $(\partial\vec{\mu}_\nu/\partial Q_\nu)$  can be nonvanishing. Several polarization directions for in- and outcoupling of resonant and up-converted fields are thus conceivable.

The calculations leading to the parametric optomechanical vacuum coupling rate  $g_{\text{opt},0}$  between the antenna field and the vibrational mode has been previously described [6], and its value is given by  $g_{\text{opt},0} = \omega_c [\vec{e}_{\text{opt}} \cdot (\partial\alpha_\nu/\partial Q_\nu) \cdot \vec{e}_{\text{opt}}] \times (1/V_{\text{opt}}\epsilon_0) \sqrt{\hbar/2\omega_\nu}$  with  $\alpha_\nu$  the polarizability tensor,  $Q_\nu$  the reduced displacement coordinate of the vibrational mode labeled by  $\nu$ ,  $V_{\text{opt}}$  the optical mode volume, and  $\vec{e}_{\text{opt}}$  the unit polarization vector of the optical antenna mode.

The coupling rate  $g_{\text{IR},0}$  associated with a vibrational mode  $\nu$  is linked to an effective transition dipole  $\vec{d}_\nu$  that can be numerically computed using, e.g., density functional theory (DFT); cf. Appendix D. Its value is  $g_{\text{IR},0} = (1/\hbar)\vec{d}_\nu \cdot \vec{E}_0$ , where the electric field per photon is given by  $\vec{E}_0 = \sqrt{\hbar\omega_\nu/2\epsilon_0 V_{\text{IR}}}\vec{e}_{\text{IR}}$  with  $V_{\text{IR}}$  the mode volume and  $\vec{e}_{\text{IR}}$  the unit polarization vector of the IR mode [27,33].

Since  $g_{\text{IR},0}^{(N_{\text{IR}})}$  scales with  $\sqrt{N_{\text{IR}}/V_{\text{IR}}}$ , it can be independent of the mode volume as long as this volume is filled with molecules. On the contrary, the interaction of the vibration with the VIS-NIR optical field  $g_{\text{opt},0}^{(N_{\text{opt}})}$  scales as  $\sqrt{N_{\text{opt}}/V_{\text{opt}}}$ , advocating for a device confining strongly this field, thereby reducing the optical power required to reach an efficient conversion process.

#### IV. DUAL PLASMONIC ANTENNA

Nanoantennas have proven to be instrumental in enhancing the interaction of molecules with off-resonant VIS-NIR optical fields [e.g., for surface-enhanced Raman scattering (SERS)] [34,35] and resonant IR fields [e.g., for surface-enhanced infrared absorption (SEIRA)] [36,37]. We now present the design of a new dual-resonant antenna (see Figs. 1 and 2) and compute the interaction of the local fields at the VIS-NIR and IR resonances with a specific vibrational mode of an ensemble of molecules covering the nanostructure. We assume that molecules are attached with their main axis perpendicular to the metallic surfaces and extract from DFT calculations the relevant components of the derivatives of the electronic moment and polarizability with respect to the normal mode coordinate. We note that calculations for specific self-assembled monolayer orientations [38,39] or randomly oriented molecules can be achieved from the full knowledge of  $(\partial\alpha_\nu/\partial Q_\nu)$  and  $(\partial\vec{\mu}_\nu/\partial Q_\nu)$ .

In our design, the incoming field (to be up-converted) and the pump laser field are each resonant with a different component of the antennas arranged in a crossed configuration (cf. Appendix E for additional information on the design and values of its parameters). At their intersection, the near-field polarizations of the two fields are collinear ( $\vec{e}_{\text{IR}} \simeq \vec{e}_{\text{opt}}$ ), and we obtain  $\eta_{\text{pol}} = 33\%$  for the vibrational mode illustrated in Fig. 2, when accounting for the specificities of the corresponding Raman tensor and IR transition dipole. Electromagnetic simulations demonstrate that the two fields, despite differing in frequency by more than one order of magnitude in that particular example, are confined within a very similar volume inside the nanogaps separating the two structures. This confinement results in a spatial overlap of the two main electromagnetic field components within the dual antenna reaching  $\eta_{\text{mode}} = 44\%$ .

From our numerical calculations, we find that the antenna-assisted IR coupling rate for the vibrational mode at wave number  $\nu = 1002 \text{ cm}^{-1}$  reaches  $g_{\text{IR},0}^{(N_{\text{IR}})}/(2\pi) \sim 186 \text{ GHz}$  as  $V_{\text{IR}}$  is decreased by several orders of magnitude below its diffraction limit (the calculation of  $V_{\text{IR}}$  and  $\vec{d}_\nu$  are detailed in Appendix C). As the cavity damping rate remains larger than the collective vacuum IR coupling rate ( $2g_{\text{IR},0}^{(N_{\text{IR}})} < \kappa^{\text{IR}}/2$ —Purcell regime), the antenna-enhanced damping rate for this vibrational mode

can be approximated by the expression [33]  $\Gamma_\nu^* \simeq \Gamma_\nu + \kappa^{\text{IR}}/2[1 - \sqrt{1 - (2g_{\text{IR},0}^{(\text{NIR})})^2/(\kappa^{\text{IR}}/2)^2}]$ .

We compute in Appendix D the coupling rate of another vibrational mode that has a larger IR dipole moment: Under optimal molecular orientation and filling conditions, that mode is at the onset of the collective strong coupling regime with the IR antenna mode [23]. While future work is needed to properly describe this regime in the context of wavelength conversion, we note that our design offers new perspectives to realize a source of IR photons. Indeed, in the strong coupling regime and under optomechanical parametric amplification [6], the optically pumped population is shared between the collective vibrational mode and the corresponding IR antenna mode, resulting in the fast emission of IR radiation.

## V. OPTICAL NOISE CONTRIBUTIONS

A useful figure of merit to compare the performance of detectors independently of their respective operational bandwidth is the noise-equivalent power:  $\text{NEP} = P_{\text{IR}}^{\text{min,in}}/\sqrt{\text{BW}}[\text{W} \cdot \text{Hz}^{-1/2}]$ , where  $P_{\text{IR}}^{\text{min,in}}$  is the incoming power at which the detector reaches a unity SNR. This definition can be translated to our converter by defining the SNR as  $\text{SNR}(\omega) = \tilde{S}_{\text{IR} \rightarrow \text{opt}}^{\text{out}}(\omega)/\tilde{S}_{\text{opt}}^{\text{out}}(\omega)$ . If we assume that the dark noise of a detector in the VIS-NIR range is negligible compared to the upconversion noise, the NEP of the converter is by extension that of a detector built upon it. We show the computed NEP in Fig. 3 as a function of the optical pump power and laser detuning from the cavity resonance. Remarkably, the NEP reaches values that improve on the state of the art for uncooled commercial devices (Table I) and compares favorably with more recently demonstrated room-temperature platforms [40,41] operating at the higher end of the frequency range achievable with our molecular up-converter.

When operating with a pump laser red-detuned from the optical resonance ( $\Delta = \omega_p - \omega_c = -\omega_\nu$ ) in the resolved sideband regime  $\kappa^{\text{opt}}/2 < \omega_\nu$ , we can simplify the interaction of Eq. (3) and obtain

$$\hat{H}_{\text{eff}} = -\hbar g_{\text{opt},0}^{(N_{\text{opt}})} \sqrt{\bar{n}_{\text{opt}}} (\delta \hat{a}_{\text{opt}}^\dagger \hat{b}_\nu + \text{H.c.}). \quad (7)$$

This regime provides maximal efficiency and optimal NEP, as seen in Fig. 3(a). We note that for low-frequency vibrational modes the condition  $\kappa^{\text{opt}}/2 < \omega_\nu$  could be achieved with the help of hybrid cavities that feature narrower linewidths [42,43].

Keeping this optimal detuning, we investigate in Fig. 3(b) how the NEP depends on optical pump power. As the intracavity pump photon number is increased, the efficiency initially grows linearly, while the noise remains constant, limited by the thermally generated anti-Stokes signal. This result yields a square-root decrease of NEP

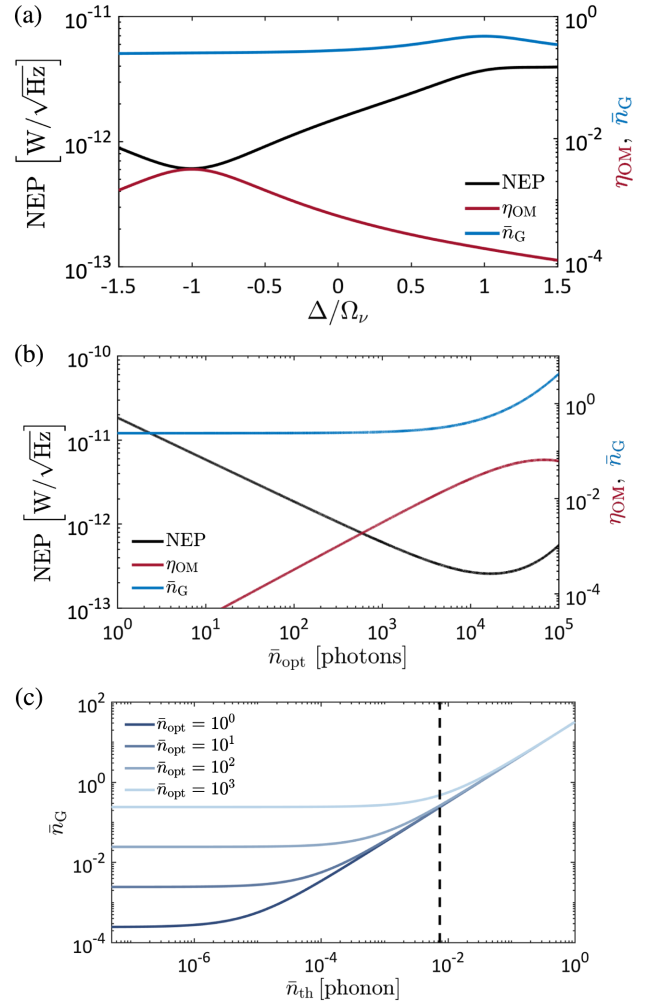


FIG. 3. (a),(b) Left axis: Noise-equivalent power (NEP) (black solid line); right axis: Power-dependent part of the internal conversion efficiency  $\eta_{\text{OM}}(\bar{n}_{\text{opt}})$  (red solid line) and noise-equivalent photon number per gate  $\bar{n}_G$  (blue solid line) plotted on a logarithmic vertical scale (a) as a function of the detuning of the optical pump laser with respect to the plasmonic resonance for a fixed intracavity optical photon number  $\bar{n}_{\text{opt}} = 1000$  and (b) plotted as a function of intracavity optical photon number  $\bar{n}_{\text{opt}}$  for a fixed optimal red detuning  $\Delta = -\Omega_\nu$ . We use the parameters for the dual antenna and molecular system described in the text. All their numerical values can be found in the Appendixes. (c)  $\bar{n}_G$  as a function of the thermal occupancy of the vibrational mode for several intracavity optical photon numbers. The dashed vertical line denotes the thermal occupancy of the mode considered in the text  $\omega_\nu/(2\pi) = 30$  THz at room temperature.

with pump power. Interestingly, at a high intracavity photon number, the contribution of optomechanical quantum back-action to the dark-count rate surpasses the thermal contribution, and the NEP degrades with increasing power. This behavior is reminiscent of the standard quantum limit for displacement detection in optomechanical cavities.

Another unique feature of the upconversion scheme is its compatibility with single-photon detectors already

TABLE I. Noise-equivalent power of commercially available uncooled devices in the 5–50 THz region [4,5] and comparison with the device presented in this manuscript (molecular device).

Detector type	NEP $\mathcal{O}(\dots)[\text{W} \cdot \text{Hz}^{-1/2}]$
Golay cell	$\mathcal{O}(10^{-9})$
Pyroelectric	$\mathcal{O}(10^{-10})$
MCT photodiode	$\mathcal{O}(10^{-10})$
Microbolometer	$\mathcal{O}(10^{-11})$
Molecular device	$\mathcal{O}(10^{-(11\dots 12)})$

available in the VIS-NIR range. To assess more precisely the feasibility of operating our device in single-photon counting mode, we introduce the noise-equivalent photon rate, i.e.,  $\tilde{S}_{\text{opt}}^{\text{out}}/\eta_{\text{ext}} \equiv |\langle \hat{a}_{\text{IR}}^{\text{in}} \rangle|^2/\text{SNR}$ . This quantity corresponds to the incoming IR photon rate at the input of the device that would generate an output rate of up-converted photons equal to the dark-count rate.

In practice, noisy single-photon detectors are best operated in gated mode. In our approach, this mode is easily realized using a pulsed optical pump laser, with a pulse duration  $\Delta t$  of a few picoseconds that ideally matches the molecular vibrational linewidth;  $\Delta t \simeq (\Gamma_{\nu}^*)^{-1}$ . This mode of operation provides not only better noise rejection and higher intracavity photon numbers (therefore better efficiency), but also ultrafast timing resolution that is not otherwise achievable due to the intrinsic timing jitter of VIS-NIR single-photon counters (typically several tens of picoseconds) [8]. We therefore define the noise-equivalent photon number per gate  $\bar{n}_G = \tilde{S}_{\text{opt}}^{\text{out}}/(\eta_{\text{ext}}\Gamma_{\nu}^*)$ , which translates the noise-equivalent photon rate into an average incoming IR noise photons per time gate.

## VI. CONVERTER ARRAYS

To gain more insight on the limiting factors constraining single-photon operation, we plot  $\bar{n}_G$  as a function of the thermal occupancy of the vibration and of the intracavity pump photon number in Fig. 3(c). This graph shows that moderate cooling of the device to 100 °C below ambient (achieved with thermoelectric cooling systems) would bring the thermal occupancy of this vibrational mode down to  $\bar{n}_{\text{th}} = 5.6 \times 10^{-4}$ , allowing one to reach a noise as low as  $\bar{n}_G \simeq 2 \times 10^{-2}$ , making single-photon counting with picosecond time resolution in the MIR and FIR domains a realistic prospect. For the lower-frequency range (5–20 THz), reducing the temperature of the molecules in a cryogenic environment would be required to allow single-photon operation.

A promising way to further reduce the gated dark-count level consists in designing an array of molecular converters, sufficiently distant from each other so as not to interact by near-field coupling. We assume that the array is illuminated by a spatially coherent IR signal and optical pump beam, which is achievable when using a high  $f$ -number lens

due to the subwavelength dimensions of the antennas. The key advantage of this scheme is that the anti-Stokes fields of thermal origin from different antennas would not exhibit any mutual phase coherence; they will add up incoherently in the far field. On the contrary, the up-converted (sum-frequency) anti-Stokes fields would be phase coherent and interfere constructively in specific directions, in analogy with a phased emitter array [44,45]. Considering a simple linear array, as demonstrated in Appendix F, this effect would jointly decrease the thermal contribution to the dark-count rate and dilute the intracavity photon number per device, enabling single-photon operation with improved sensitivity.

A configuration with multiple converters within the IR spot could alternatively be leveraged for on-chip IR multiplexing [46–49] with distinct converters responding to distinct IR frequencies by the proper choice of molecular vibrations and antenna design, thereby bypassing the limited detection bandwidth of a single converter. This subwavelength platform benefits from the coherent nature of the conversion process and opens the route to IR spectroscopy, IR hyperspectral imaging, and recognition technologies.

## VII. CONCLUSION

In summary, we presented a new concept for frequency upconversion from the mid-infrared to the visible domain based on the interaction of both fields with molecular vibrations coupled to a dual-resonant nanoantenna. We considered an incoming long-wavelength infrared radiation that resonantly excites a vibrational mode, which is simultaneously coupled through its Raman polarizability to a coherent pump field at a shorter wavelength (visible or near-infrared), resulting in upconversion of the IR signal onto the anti-Stokes sideband of the pump. Thanks to the recently developed framework of molecular cavity optomechanics, we were able to treat the problem in a full quantum model and, thereby, predict the internal quantum efficiency of our device, as well as its outgoing noise spectral density. We showed that the noise added in the conversion process has two main origins: thermal noise and backaction noise (including quantum and dynamical backaction), the latter increasing superlinearly with pump power and eventually becoming dominant. We analyzed the dependence of the NEP on the intracavity pump photon number and pump-cavity detuning and predicted that, under the optimal condition of red-sideband excitation, the NEP can be as low as a few  $\text{pW} \cdot \text{Hz}^{-1/2}$ , improving on the state of the art for devices operating at ambient conditions.

We stress that our numerical estimates are based on a realistic nanoantenna design and a common simple molecule (thiophenol). Although the intracavity photon numbers required to reach optimal performance appear to be large, they can be achieved under pulsed excitation [50].

Moreover, requirements on the intracavity power would be lowered by further reducing the gap size (down to 1–2 nm) and by chemical engineering of the molecular converter toward higher Raman activity. Our study also shows that, by moderately increasing the resonant coupling rate between molecular vibration and IR antenna, the system would enter the IR strong coupling regime, with the formation of vibrational polaritons [23]. We leave the study of the conversion process in this new regime for future investigation.

Code to reproduce the data in Figs. 3 and 4 in addition to the antenna and molecular parameters is available on Zenodo [51].

## ACKNOWLEDGMENTS

The authors thank Wen Chen and the reviewers of the article for valuable comments. C. G. acknowledges support from the Swiss National Science Foundation through Grant No. PP00P2-170684. This work has received funding from the European Union's Horizon 2020 research and innovation program with Grants Agreement No. 732894 (HOT) and No. 829067 (THOR). P. R. acknowledges support from the Max Planck-EPFL Center for Molecular Nanoscience and Technology. D. M.-C. thanks Vahid Sandoghdar and acknowledges financial support from the Max Planck Society.

## APPENDIX A: OPTOMECHANICAL FRAMEWORK

A detailed description of the optomechanical framework can be found elsewhere [6,28–30]. Here, we just remind the readers of the few definitions and relationships used in the paper.

The average number of intracavity excitations in the VIS-NIR (label “opt”) or IR antenna mode is related to the incoming photon flux  $|\langle \hat{a}_{\text{opt/IR}}^{\text{in}} \rangle|$  and incoming power  $P_{\text{opt/IR}}^{\text{in}}$  by

$$\begin{aligned} \bar{n}_{\text{opt/IR}} &= |\langle \hat{a}_{\text{opt/IR}}^{\text{in}} \rangle|^2 \frac{\kappa_{\text{ex}}^{\text{opt/IR}}}{\Delta^2 + (\kappa^{\text{opt/IR}}/2)^2} \\ &= \frac{P_{\text{opt/IR}}^{\text{in}}}{\hbar\omega_{p/\text{IR}}} \frac{\kappa_{\text{ex}}^{\text{opt/IR}}}{\Delta^2 + (\kappa^{\text{opt/IR}}/2)^2}. \end{aligned} \quad (\text{A1})$$

When considering the molecular vibrational levels and their parametric coupling to the optical field, the antenna-assisted transition rate to a lower excited level (anti-Stokes transition) is given by

$$A^- = (g_{\text{opt},0}^{(N_{\text{opt}})})^2 \bar{n}_{\text{opt}} \frac{\kappa^{\text{opt}}}{(\omega_{\nu} - \Delta)^2 + (\kappa^{\text{opt}}/2)^2}. \quad (\text{A2})$$

The antenna-assisted transition rate to a higher excited vibrational level (Stokes transition) is given by

$$A^+ = (g_{\text{opt},0}^{(N_{\text{opt}})})^2 \bar{n}_{\text{opt}} \frac{\kappa^{\text{opt}}}{(\omega_{\nu} + \Delta)^2 + (\kappa^{\text{opt}}/2)^2}. \quad (\text{A3})$$

The interested reader can find the complete derivation of the outgoing spectral density in Ref. [26]. In this manuscript, we are interested in the signal arising on the anti-Stokes sideband. Starting from Eq. (4) in the main text and following the same calculation steps, we arrive at the final expression

$$S_{\text{tot}}^{\text{out}}(\omega_{\text{aS}}) = \frac{2}{\pi} \frac{\eta_{\text{opt}} A^-}{\Gamma_{\nu}^* + \Gamma_{\text{opt}}} \bar{n}_f. \quad (\text{A4})$$

For convenience, we label the different components of the outgoing noise spectral density according to the origin of the vibrational population from which they result [cf. Eq. (5) in the main text]:

$$\begin{aligned} S_{\text{tot}}^{\text{out}}(\omega_{\text{aS}}) &\propto \underbrace{\frac{A^-}{\Gamma_{\nu}^* + \Gamma_{\text{opt}}} \bar{n}_{\text{th}}}_{S_{\text{th}}^{\text{out}}} + \underbrace{\frac{A^- \Gamma_{\nu}^*}{(\Gamma_{\nu}^* + \Gamma_{\text{opt}})^2} \eta_{\text{overlap}} \bar{n}_b^{\text{IR}}}_{S_{\text{IR} \rightarrow \text{opt}}^{\text{out}}} \\ &\quad + \underbrace{\frac{A^-}{(\Gamma_{\nu}^* + \Gamma_{\text{opt}})^2} [A^+ - \Gamma_{\text{opt}} \bar{n}_{\text{th}}]}_{S_{\text{ba}}^{\text{out}}}. \end{aligned} \quad (\text{A5})$$

With this notation, the total noise quanta in the outgoing optical field is  $S_{\text{opt}}^{\text{out}} = S_{\text{th}}^{\text{out}} + S_{\text{ba}}^{\text{out}}$ .

We can then derive the expression for the conversion efficiency defined as  $\tilde{S}_{\text{IR} \rightarrow \text{opt}}^{\text{out}} = \eta_{\text{ext}} |\langle \hat{a}_{\text{IR}}^{\text{in}} \rangle|^2$  and obtain

$$\eta_{\text{ext}} = \eta_{\text{opt}} \cdot \eta_{\text{overlap}} \cdot \underbrace{\frac{A^- \Gamma_{\nu}^*}{\Gamma_{\nu}^* + \Gamma_{\text{opt}}} \frac{1}{\kappa^{\text{IR}}} \left( \frac{2g_{\text{IR},0}^{(N_{\text{IR}})}}{\Gamma_{\nu}^* + \Gamma_{\text{opt}}} \right)^2}_{\eta_{\text{int}}} \cdot \eta_{\text{IR}}. \quad (\text{A6})$$

This expression comprises the different factors constituting Eq. (6) of the main text.

For a pump field red-detuned from the optical antenna resonance ( $\Delta = \omega_p - \omega_c = -\omega_{\nu}$ ), this expression can be further developed to evidence the dependence of the internal conversion efficiency on IR and optical collective vacuum coupling rates:

$$\eta_{\text{int}} = \eta_{\text{overlap}} \cdot \frac{(2g_{\text{opt},0}^{(N_{\text{opt}})})^2 \bar{n}_{\text{opt}}}{\kappa^{\text{opt}} (\Gamma_{\nu}^* + \Gamma_{\text{opt}})} \frac{\Gamma_{\nu}^*}{(\Gamma_{\nu}^* + \Gamma_{\text{opt}})} \frac{(2g_{\text{IR},0}^{(N_{\text{IR}})})^2}{\kappa^{\text{IR}} (\Gamma_{\nu}^* + \Gamma_{\text{opt}})}. \quad (\text{A7})$$

Using this conversion efficiency, we have an alternate way to calculate the NEP directly from the dark-count rate and efficiency of the device as  $\text{NEP} = (\hbar\omega_{\nu}/\eta_{\text{ext}}) \sqrt{\tilde{S}_{\text{opt}}^{\text{out}}}$  [8].

This method gives identical results as the one presented in the main text.



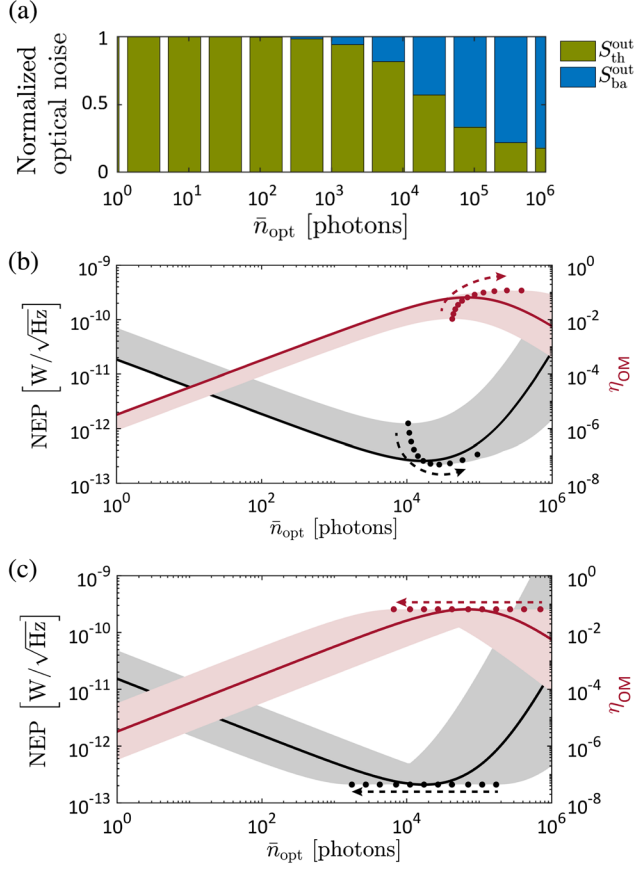


FIG. 4. (a) Relative contributions to outgoing optical noise ( $S_{\text{opt}}^{\text{out}}$ ) on the anti-Stokes sideband as a function of the intracavity photon number for a pump tone red-detuned from the resonance of the optical antenna ( $\Delta = -\omega_\nu$ ). The contribution from the thermal population of the vibrational mode to the dark-count rate is depicted in green and the backaction noise in blue. (b),(c) NEP (left axis, black solid line) and power-dependent part of the internal conversion efficiency ( $\eta_{\text{OM}}$ , right axis, red solid line) as a function of the intracavity optical photon number for  $\Delta = -\omega_\nu$ . The parameters used to plot the lines are described in the Appendixes. The dots indicate the extremal values in NEP and  $\eta_{\text{OM}}$  when varying the absorption intensity  $[0.1:10]I_\nu^{\text{IR}}$  in (b) or the Raman activity  $[0.1:10]R_\nu^{\text{LL}}$  in (c). The colored areas denote the NEP and  $\eta_{\text{OM}}$  achievable when sweeping either parameter, and the dashed arrows indicate the direction of evolution of the extremum when increasing the parameter value.

In Fig. 4, we show the computed conversion efficiency and NEP for the case  $\Delta = -\omega_\nu$ . In this red-detuned configuration, our model assumptions remain valid for a large range of optical intracavity photon numbers. At high optical power, we observe that the efficiency and the NEP reach an extremal value when the backaction contribution to the outgoing noise equals that of thermal vibrations.

In Figs. 4(b) and 4(c), we also show how these extremal points move when changing the IR absorption cross section [in Fig. 4(b)] or Raman activity [in Fig. 4(c)] of the molecules. As can be expected, improving the Raman

activity of the molecules only displaces the curves to the left without modifying the extremal values—meaning that lower optical pump power is required to reach the same NEP and conversion efficiency. In contrast, improving the IR absorption can lead to a lower achievable NEP, in a certain range. For too large IR cross sections, the vibration starts to enter the strong coupling regime with the IR antenna, and a different treatment is needed to provide accurate predictions.

## APPENDIX B: ZERO-TEMPERATURE LIMIT

The limit of vanishing thermal occupancy of the vibrational mode is relevant for specific applications, and it demonstrates how the backaction noise acting onto the vibration sets a fundamental lower bound on the achievable NEP of the converter. When  $\bar{n}_{\text{th}} \sim 0$ , the outgoing noise spectral density of Eq. (A5) can be simplified to

$$S_{0\text{K}}^{\text{out}}(\omega_{\text{aS}}) = \frac{2}{\pi} \eta_{\text{opt}} \frac{A^- A^+}{(\Gamma_\nu^* + \Gamma_{\text{opt}})^2}. \quad (\text{B1})$$

Taking into account Eq. (A6) for the expression of the external conversion efficiency, the NEP and the  $\bar{n}_G$  at 0 K can be calculated.

In the regime of weak optical pumping,  $\Gamma_{\text{opt}} \ll \Gamma_\nu^*$ , we obtain a linear scaling of  $\bar{n}_G$  as a function of the intracavity photon number (appearing in the transition rates  $A^-$  and  $A^+$ ) while the NEP remains constant (cf. Fig. 5). The value of the NEP at this plateau, which corresponds to an intrinsic quantum limit due to measurement backaction, is given by the following expression:

$$\text{NEP}_{0\text{K}} = \frac{\hbar \omega_\nu \kappa^{\text{IR}}}{\eta_{\text{overlap}} \sqrt{\eta_{\text{opt}} \eta_{\text{IR}}}} \sqrt{\frac{A^+ (\Gamma_\nu^* + \Gamma_{\text{opt}})^{5/2}}{A^- \Gamma_\nu^* (2g_{\text{IR},0}^{(N_{\text{IR}})})^2}}. \quad (\text{B2})$$

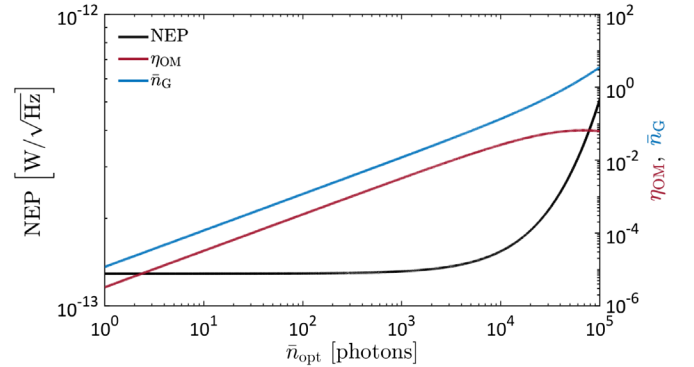


FIG. 5. Zero-temperature quantum limit on the NEP. NEP (black solid line, left axis), power-dependent part of the internal conversion efficiency ( $\eta_{\text{OM}}$ , red solid line, right axis), and noise-equivalent photon number per gate ( $\bar{n}_G$ , blue solid line, right axis) as a function of the intracavity optical photon number for  $\Delta = -\omega_\nu$ . The parameters used here are described in the Appendixes.



### APPENDIX C: ABSORPTION OF INCOMING IR RADIATION BY MOLECULAR VIBRATION

We describe in the following the coupling between a resonant field and a single molecular oscillator inside a cavity (i.e., antenna). We also derive the expression given in the main text for the number of excited phonons in the steady state. Our treatment is inspired by that of Ref. [24].

The interaction between an external monochromatic field of frequency  $\omega_{\text{IR}}$  and the molecular vibration inside a cavity is given in the dipole approximation by

$$\hat{H}_{\text{int}} = -\vec{d} \cdot \vec{\mathcal{E}}_{\text{IR}}, \quad (\text{C1})$$

where  $\vec{\mathcal{E}}_{\text{IR}} = i\sqrt{\bar{n}_{\text{IR}}}[e^{-i\omega_{\text{IR}}t}e^{-i(\phi+\phi_0)t}\vec{\mathcal{E}}_0 - e^{i\omega_{\text{IR}}t}e^{i(\phi+\phi_0)t}\vec{\mathcal{E}}_0^*]$  with  $\phi$  the phase offset between the field and the dipole and  $\phi_0$  an adjustable phase parameter of the driving field.  $\vec{\mathcal{E}}_0 = \sqrt{\hbar\omega_{\text{IR}}/2\epsilon_0V_{\text{IR}}}\vec{e}_k$  is the vacuum field, with  $V_{\text{IR}}$  the mode volume and  $\vec{e}_k$  the unit polarization vector of the IR mode.

This interaction can be written in terms of the bosonic ladder operators describing the IR mode inside the cavity  $\hat{a}_{\text{IR}}^\dagger$  and  $\hat{a}_{\text{IR}}$  and the vibrational phononic operators  $\hat{b}_\nu^\dagger$  and  $\hat{b}_\nu$  at frequency  $\nu$ . For a weak IR drive, the vibrational Hilbert space of each molecule can be reduced to ground and first excited state  $\{|0\rangle, |1\rangle\}$  and described like a two-level system (TLS) with creation and annihilation operators  $\hat{\sigma}_\nu^+$  and  $\hat{\sigma}_\nu^-$  [27]. We note that the validity of the TLS description for a collective vibrational mode of  $N$  oscillators would break down only for a number of excitations of the order of  $N$  [33].

The dipolar transition is purely off-diagonal in this basis and described as  $\vec{d} = d_\nu(\hat{\sigma}_\nu^- \vec{e}_\nu + \hat{\sigma}_\nu^+ \vec{e}_\nu)$ . The field inside the cavity is, in turn, described by  $\vec{\mathcal{E}}_{\text{IR}} = -i\sqrt{\bar{n}_{\text{IR}}}\mathcal{E}_0[e^{-i\phi_0}\hat{a}_{\text{IR}}\vec{e}_k - e^{i\phi_0}\hat{a}_{\text{IR}}^\dagger\vec{e}_k^*]$  [33]. In a frame rotating at the frequency of the IR driving field and keeping only the resonant processes ( $\hat{a}_{\text{IR}}^\dagger\hat{\sigma}_\nu^-$ ,  $\hat{a}_{\text{IR}}\hat{\sigma}_\nu^+$ ), we obtain the interaction Hamiltonian:

$$\hat{H}_{\text{int}} = -i\hbar g_{\text{IR}}(e^{-i\phi_0}\hat{a}_{\text{IR}}^\dagger\hat{\sigma}_\nu^- + e^{i\phi_0}\hat{a}_{\text{IR}}\hat{\sigma}_\nu^+), \quad (\text{C2})$$

with  $g_{\text{IR}} = (d_\nu \cdot \mathcal{E}_0/\hbar)\sqrt{\bar{n}_{\text{IR}}}\vec{e}_\nu\vec{e}_k e^{i\phi_0} = g_{\text{IR},0}\sqrt{\bar{n}_{\text{IR}}}$ .

Here, we choose the additional phase term of the driving field in order for the coupling to be real positive, without loss of generality.

We follow the dynamics of the TLS in this rotating frame. Introducing the rate  $\Gamma_{\text{tot}}$ , which describes the total damping of the vibrational mode as described in the main text, we obtain

$$\dot{\rho}_{11} = i\frac{g_{\text{IR}}}{2}(\rho_{10} - \rho_{01}) - \Gamma_{\text{tot}}\rho_{11}, \quad (\text{C3a})$$

$$\dot{\rho}_{00} = -i\frac{g_{\text{IR}}}{2}(\rho_{10} - \rho_{01}) + \Gamma_{\text{tot}}\rho_{11}, \quad (\text{C3b})$$

$$\dot{\rho}_{01} = -i\delta\rho_{01} - i\frac{g_{\text{IR}}}{2}(\rho_{11} - \rho_{00}) - \frac{\Gamma_{\text{tot}}}{2}\rho_{01}, \quad (\text{C3c})$$

$$\dot{\rho}_{10} = i\delta\rho_{10} + i\frac{g_{\text{IR}}}{2}(\rho_{11} - \rho_{00}) - \frac{\Gamma_{\text{tot}}}{2}\rho_{10}, \quad (\text{C3d})$$

with  $\delta = \omega_{\text{IR}} - \omega_\nu$  the detuning between the IR drive and the vibrational resonance.

These equations are often described with the help of the Bloch vector components:

$$u = \frac{1}{2}(\rho_{01} + \rho_{10}), \quad (\text{C4a})$$

$$v = \frac{1}{2i}(\rho_{01} - \rho_{10}), \quad (\text{C4b})$$

$$w = \frac{1}{2}(\rho_{11} - \rho_{00}). \quad (\text{C4c})$$

The components  $u$  and  $v$  of the Bloch vectors are related to the average dipole value [24]:  $\langle \vec{d} \rangle = 2\vec{d}_\nu(u \cos \omega_{\text{IR}}t - v \sin \omega_{\text{IR}}t)$ . We derive the master equations as a function of these components:

$$\dot{u} = \delta v - \frac{\Gamma_{\text{tot}}}{2}u, \quad (\text{C5a})$$

$$\dot{v} = -\delta u - g_{\text{IR}}w - \frac{\Gamma_{\text{tot}}}{2}v, \quad (\text{C5b})$$

$$\dot{w} = g_{\text{IR}}v - \Gamma_{\text{tot}}w - \frac{\Gamma_{\text{tot}}}{2}. \quad (\text{C5c})$$

The steady-state solutions of these equations are

$$\bar{u} = \frac{g_{\text{IR}}}{2} \frac{\delta}{\delta^2 + (\Gamma_{\text{tot}}^2/4) + (g_{\text{IR}}^2/2)}, \quad (\text{C6a})$$

$$\bar{v} = \frac{g_{\text{IR}}}{2} \frac{\Gamma/2}{\delta^2 + (\Gamma_{\text{tot}}^2/4) + (g_{\text{IR}}^2/2)}, \quad (\text{C6b})$$

$$\bar{w} + \frac{1}{2} = \frac{g_{\text{IR}}^2}{4} \frac{1}{\delta^2 + (\Gamma_{\text{tot}}^2/4) + (g_{\text{IR}}^2/2)}. \quad (\text{C6c})$$

The average number of photons absorbed per unit time by the vibrational dipole is given by

$$\frac{d\bar{n}_b^{\text{IR}}}{dt} = \frac{d\bar{W}^{\text{IR}}}{dt} \frac{1}{\hbar\omega_{\text{IR}}} = \frac{\vec{\mathcal{E}}_n \cos \omega_{\text{IR}}t \cdot \langle \dot{\vec{d}} \rangle}{\hbar\omega_{\text{IR}}}. \quad (\text{C7})$$

If the detuning and coupling are much smaller than the vibrational damping rate ( $\delta, g_{\text{IR}} < \Gamma_{\text{tot}}$ ), the average number of absorptions over an IR period can be written as

$$\frac{d\bar{n}_b^{\text{IR}}}{dt} = g_{\text{IR}} \bar{v} = \frac{g_{\text{IR}}^2}{2} \frac{\Gamma_{\text{tot}}/2}{\delta^2 + (\Gamma_{\text{tot}}^2/4) + (g_{\text{IR}}^2/2)} \approx \frac{g_{\text{IR}}^2}{\Gamma_{\text{tot}}}. \quad (\text{C8})$$

In the steady state, the rate of photons absorbed by the vibrational mode equals the phonon damping rate so that the average number of excited phonons is

$$\bar{n}_b^{\text{IR}} = \frac{d\bar{n}_b^{\text{IR}}}{dt} \frac{1}{\Gamma_{\text{tot}}} = \frac{g_{\text{IR}}^2}{\Gamma_{\text{tot}}^2} = \frac{g_{\text{IR},0}^2}{\Gamma_{\text{tot}}^2} \bar{n}_{\text{IR}} = \frac{g_{\text{IR},0}^2}{\Gamma_{\text{tot}}^2} \frac{\kappa_{\text{ex}}^{\text{IR}}}{\delta^2 + (\kappa_{\text{IR}}^{\text{IR}}/2)^2} \tilde{S}_{\text{IR}}^{\text{in}} \stackrel{\delta \ll \kappa_{\text{IR}}^{\text{IR}}}{\approx} \frac{4g_{\text{IR},0}^2}{\Gamma_{\text{tot}}^2} \frac{\eta_{\text{IR}}}{\kappa_{\text{IR}}^{\text{IR}}} |\langle \hat{a}_{\text{IR}}^{\text{in}} \rangle|^2. \quad (\text{C9})$$

The condition  $\delta \ll \kappa_{\text{IR}}^{\text{IR}}$  is satisfied in realistic scenarios, since the IR antenna decay rate  $\kappa_{\text{IR}}^{\text{IR}}$  is typically faster than the vibrational damping rate  $\Gamma_{\text{tot}}$ . We note that the average number of excited phonons  $\bar{n}_b^{\text{IR}}$  can also be simply derived from the steady-state population of the upper TLS state  $\bar{n}_b^{\text{IR}} = \bar{w} + \frac{1}{2}$ .

## APPENDIX D: SIMULATION OF MOLECULAR PARAMETERS

The infrared absorption intensity of fundamental vibrational transitions  $I_\nu^{\text{IR}}$  can be obtained by DFT calculations [32,52] of the derivatives of the electric moment components  $\mu_\nu^i$  with respect to the normal coordinates representing the vibrational mode of interest. They are usually expressed in [ $\text{km} \cdot \text{mol}^{-1}$ ]. For a nondegenerate and harmonic vibrational mode, the absorption intensity averaged over all orientations is given by

$$\langle I_\nu^{\text{IR}} \rangle = \frac{N_A}{12\epsilon_0 c^2} \sum_{i=1}^3 \left( \frac{\partial \mu_\nu^i}{\partial Q_\nu} \right)^2, \quad (\text{D1})$$

with  $N_A$  the Avogadro number.

### 1. GAUSSIAN calculations

The procedure is well described in the context of Raman calculations in the book of Le Ru and Etchegoin [53]. The GAUSSIAN software gives access to the derivatives of the electric dipole with respect to the  $i$ th component of the displacement in Cartesian coordinates of the  $n$ th atom. These derivatives can then be converted to derivatives with respect to the normal coordinates of a vibrational mode  $\nu$  and are given in atomic units; i.e., the electric moment is given in bohr-electron ( $2.54 \text{ D}/8.48 \times 10^{-30} \text{ C} \cdot \text{m}$ ) and the displacement in bohr ( $0.529 \text{ \AA}$ ). The quantities  $\partial \vec{\mu}_\nu / \partial Q_\nu$  can be converted to other systems of units:

$$\begin{aligned} & \left( \frac{\partial \vec{\mu}_\nu}{\partial Q_\nu} \right)^2 [\text{D}^2 \cdot \text{\AA}^{-2} \cdot \text{amu}^{-1}] \\ &= \left[ \frac{2.54}{0.53} \right]^2 \cdot \left( \frac{\partial \vec{\mu}_\nu}{\partial Q_\nu} \right)^2 [\text{atomic units}] \end{aligned} \quad (\text{D2})$$

and, finally, linked to the absorption intensity  $I_\nu^{\text{IR}}$  of an incoming field of polarization  $\vec{e}_i$ :

$$I_\nu^{\text{IR}} [\text{km} \cdot \text{mol}^{-1}] = 126.8 \cdot \left( \vec{e}_i \cdot \frac{\partial \vec{\mu}_\nu}{\partial Q_\nu} \right)^2 [\text{D}^2 \cdot \text{\AA}^{-2} \cdot \text{amu}^{-1}]. \quad (\text{D3})$$

## 2. Cross section

This expression can be linked to the absorption cross section  $I_\nu^{\text{IR}} = N_A \int \sigma_{\nu,\text{abs}} d\nu'$ . If we assume a Lorentzian profile for the transition considered, the on-resonance value of the cross section is  $\int \sigma_{\nu,\text{abs}} d\nu' = (\pi/2) \sigma_{\text{abs}}(\nu' = \nu) \delta\nu = (\pi/2) \sigma_{\text{abs}}(\nu' = \nu) (\Gamma_{\text{tot}}/2\pi c)$ .

Thus, the absorption cross section can be inferred from DFT calculations:

$$\sigma_{\nu,\text{abs}} [\text{cm}^2] = \frac{4c}{N_A \Gamma_{\text{tot}}} I_\nu^{\text{IR}} [\text{km} \cdot \text{mol}^{-1}] \times 10^7. \quad (\text{D4})$$

## 3. Effective dipole moment

Accordingly, we can also describe an effective dipole moment  $d_\nu$  to characterize the vibrational transition and link it explicitly to the electronic moment derivatives found in DFT calculations:

$$d_\nu = \sqrt{\frac{3\hbar\epsilon_0 c \Gamma_{\text{tot}} \sigma_{\nu,\text{abs}}}{2\omega_\nu}} = \sqrt{\frac{6\hbar\epsilon_0 c^2 10^3}{N_A}} \sqrt{\frac{I_\nu^{\text{IR}} [\text{km} \cdot \text{mol}^{-1}]}{\omega_\nu}}. \quad (\text{D5})$$

## 4. Raman activity of an ensemble of molecules

We refer the interested reader to Refs. [6,53] for detailed descriptions of the Raman activity, its connection with the optomechanical coupling rate, and its calculation through DFT. For completeness, we reproduce here a few expressions of the tensorial quantity  $\partial \alpha_\nu / \partial Q_\nu$  averaged over randomly oriented molecules. To simplify the notation, we introduce the Raman tensor  $R_\nu = (\partial \alpha_\nu / \partial Q_\nu)$ , and we refer to the scalar  $(\vec{e}_i \cdot R_\nu \cdot \vec{e}_j)$  as  $R_\nu^{ij}$ . Taking  $\vec{e}_i \perp \vec{e}_j$  and averaging over random orientations of the molecules, one can obtain

$$\langle |R_\nu^{ii}|^2 \rangle = \sqrt{\frac{45\bar{\alpha}_\nu^2 + 4\bar{\gamma}_\nu^2}{45}}, \quad (\text{D6})$$

$$\langle |R_\nu^{jj}|^2 \rangle = \sqrt{\frac{3\bar{\gamma}_\nu^2}{45}}, \quad (\text{D7})$$

with  $\bar{\alpha}_\nu^2 = \frac{1}{9} [R_\nu^{xx} + R_\nu^{yy} + R_\nu^{zz}]^2$  and  $\bar{\gamma}_\nu^2 = \frac{1}{2} [(R_\nu^{xx} - R_\nu^{yy})^2 + (R_\nu^{yy} - R_\nu^{zz})^2 + (R_\nu^{zz} - R_\nu^{xx})^2] + 3[(R_\nu^{xy})^2 + (R_\nu^{xz})^2 + (R_\nu^{yz})^2]$ . These quantities do not depend on the two orthogonal

orientations of the field chosen as polarization basis but only on the intrinsic properties of the molecule. In that situation, Raman scattering can be described by a scalar named the magnitude of the Raman tensor  $\mathcal{R}^2 = \langle |R_\nu^{ii}|^2 \rangle + \langle |R_\nu^{jj}|^2 \rangle = (45\bar{\alpha}_\nu^2 + 7\bar{\gamma}_\nu^2/45)$  and can be derived directly from DFT calculations.

We also introduce the depolarization ratio  $\rho = \langle |R_\nu^{jj}|^2 \rangle / \langle |R_\nu^{ii}|^2 \rangle$  that evaluates the importance of the cross-polarized component of the Raman-scattered field (with respect to the incoming field) and that is bounded by  $0 \leq \rho \leq 3/4$ . In the SERS scenario, the outgoing field is solely polarized along the direction of the local cavity field  $\vec{e}_L$ . For randomly oriented molecules, the magnitude of the Raman tensor is thus rescaled by a factor dependent on the depolarization ratio:

$$\langle |R_\nu^{LL}|^2 \rangle = \mathcal{R}^2 \frac{1}{1 + \rho}, \quad (\text{D8})$$

### 5. Local overlap— $\eta_{\text{pol}}$

The factor  $\eta_{\text{pol}}$  describes the local overlap between the two fields involved in our conversion scheme, on the one hand, and the IR dipole and Raman tensor of the molecular vibration, on the other hand. It is defined in the following way:

$$\eta_{\text{pol}} = \frac{\vec{e}_L \cdot \frac{\partial \vec{\mu}_\nu}{\partial Q_\nu} R_\nu^{LL}}{\left\| \frac{\partial \vec{\mu}_\nu}{\partial Q_\nu} \right\| \left\| R_\nu^{LL} \right\|} \quad (\text{D9})$$

with the label  $L$  designating the direction of the near field at the location of the molecule. To compute  $\langle \eta_{\text{pol}} \rangle$  (see Table II), we numerically average  $\eta_{\text{pol}}$  over all possible orientations of the molecule, while keeping the IR and optical local field collinear.

### 6. Orientation and number of molecules contributing to the IR or optical process

From the DFT calculations, we compute the molecular parameters for several cases of interest and report their values in Table II. Two orientations (main axis of the molecule parallel to both local fields or fully random molecular

orientation) are considered. Two options are also considered for the coverage: one monolayer covering the planar parts of the metallic nanostructure or a superposition of layers filling the entire volume where the fields are localized. We use the IR or optical mode volumes  $V_{\text{IR/opt}}$  (given below), the molar mass ( $M = 0.1102$  kg/mol), volume density ( $\rho = 1077$  kg/m<sup>3</sup>), or surface density ( $\rho_S = 6.8 \times 10^{18}$  m<sup>-2</sup>) of thiophenol to estimate the number  $N_{\text{IR}}$  ( $N_{\text{opt}}$ ) of molecules participating in the IR (optical) process.

## APPENDIX E: NUMERICAL SIMULATION OF THE ANTENNA'S OPTICAL RESPONSE

Our dual antenna consists of two gold bow-tie structures. We set the gap between the tips of both antennas ( $S = 25$  nm) so that the design can be fabricated using current nanofabrication techniques such as focused ion beam milling or advanced  $e$ -beam lithography. We select the other structural parameters (Fig. 6) in order to obtain appropriate resonances both in the mid-IR (length  $L$  and width  $W$ ) and in the optical domain (short length  $l$ ). In our design, the 24-nm-high nanostructure lies on top of a gold substrate. They are separated by an inactive dielectric layer ( $n = 1.47$ ) of  $8 \mu\text{m}$  thickness. This substrate reflects the incoming IR field and creates an interference pattern that improves IR absorption as shown in a previous study [36].

### 1. Numerical calculations

We use a 3D FEM software (Comsol Multiphysics) to evaluate our dual-antenna design. A Drude-Lorentz model describes the electromagnetic response of gold fitted from experimental data [54]. For the calculation in the optical range, a dielectric layer (ITO) with refractive index  $n = 1.94$  and thickness  $52$  nm is added below the antenna [55].

A dipolar emitter is placed in the center of our structure to evaluate the local density of electromagnetic states inside the antenna. Figure 6 shows the modification of the radiated power as a function of the oscillation frequency of the dipole. Based on these plots, we model the response of the structure to an incoming optical field and to an incoming mid-IR field (around 32 THz) as being each dominated by a single resonance with Lorentzian profile. We thus use a multi-Lorentzian fit to extract the relevant linewidths and

TABLE II. Molecular parameters of interest for our conversion scheme for two vibrational modes of the thiophenol molecule. Calculations are obtained for a molecule oriented vertically with respect to both IR and VIS or NIR local fields (values averaged over all molecular orientations are given in parentheses for completeness). The resulting resonant coupling terms are calculated for two different coverages of the nanostructure by the molecules and given in units of  $\kappa^{\text{IR}}$ .

Mode [cm <sup>-1</sup> ]	$I_\nu^{\text{IR}}(\langle I_\nu^{\text{IR}} \rangle)$ [km · mol <sup>-1</sup> ]	$R_\nu^{LL}(\langle R_\nu^{LL} \rangle)$ [ $\text{\AA}^4 \cdot \text{amu}^{-1}$ ]	$\eta_{\text{pol}}(\langle \eta_{\text{pol}} \rangle)$	Coverage	$g_{\text{IR},0}^{(N_{\text{IR}})} / \kappa^{\text{IR}}$
Mode 1002	0.52 (0.51)	2.40 (0.96)	0.33 (0.14)	Monolayer Volume	0.01 0.06
Mode 1093	86.95 (28.52)	0.85 (0.31)	0.97 (0.18)	Monolayer Volume	0.17 0.75

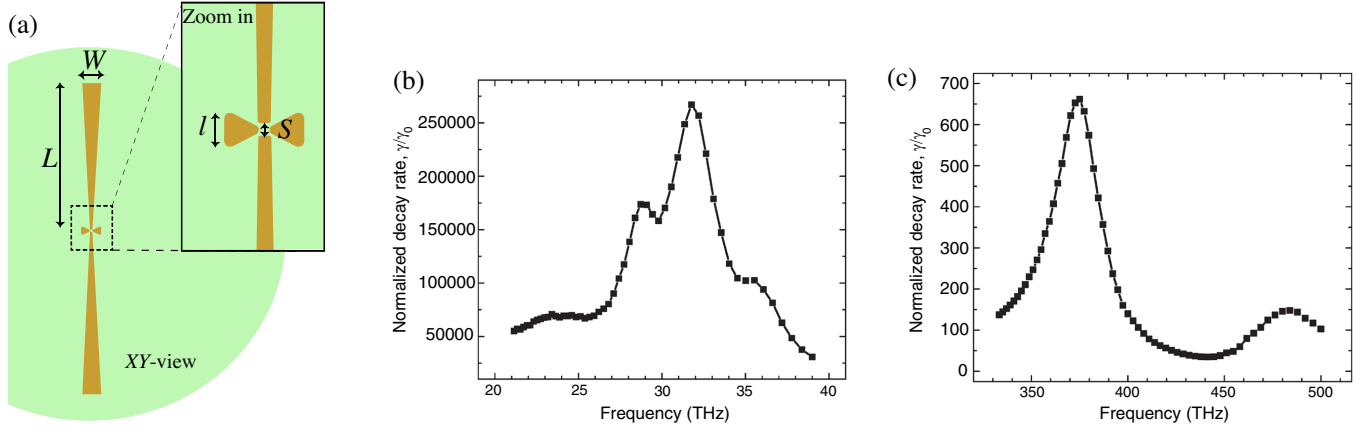


FIG. 6. (a) Geometry of the dual antenna considered in our manuscript. The parameters  $W = 70$  nm,  $L = 2.25$   $\mu$ m,  $l = 90$  nm, and  $S = 25$  nm are used to run the numerical simulations and evaluate the performance of our conversion scheme. The antenna edges are rounded with 4 and 1.3 nm radii of curvature for the optical and IR antenna, respectively. (b),(c) Normalized decay rate of an emitter placed at the center of the dual antenna in the mid-IR (b) and in the visible range (c).

total decay rates. Through the Purcell formula [30], we could estimate the corresponding effective mode volumes. Additional integrals are computed to determine the losses originating from absorption in the metal and determine the ratio between intrinsic and radiative losses at the resonance frequencies. All parameters are shown in Table III.

## 2. Spatial overlap— $\eta_{\text{mode}}$

The spatial overlap between the two electromagnetic modes is computed numerically from the field distributions of both modes according to

$$\eta_{\text{mode}} = \frac{(\int |\vec{E}_{\text{IR}} \cdot \vec{E}_{\text{opt}}| dA)^2}{(\int |\vec{E}_{\text{IR}} \cdot \vec{E}_{\text{IR}}| dA)(\int |\vec{E}_{\text{opt}} \cdot \vec{E}_{\text{opt}}| dA)}. \quad (\text{E1})$$

Imperfect overlap can result from polarization and/or confinement mismatch. In our case, the dominant mismatch is that between the spatial extents of the two modes. In the regions where both fields are confined, their polarization mismatch is, on the contrary, negligible.

TABLE III. Antenna parameters as obtained from our FEM simulations.  $\kappa^{\text{IR/opt}}$  are the total decay rates of IR (respectively, optical) energy stored in the antenna,  $\eta_{\text{IR/opt}}$  are the respective radiative efficiencies defined as  $\kappa_{\text{ex}}^{\text{IR/opt}}/\kappa^{\text{IR/opt}}$ , and  $V_{\text{IR/opt}}$  are the mode volumes.

Parameter	Mid-IR	VIS/NIR
$\kappa^{\text{IR/opt}}/2\pi$ [THz]	3.2	26.7
$\eta_{\text{IR/opt}}$	0.52	0.73
$V_{\text{IR/opt}}$ [ $\text{m}^3$ ]	$2.6 \times 10^{-21}$	$1.0 \times 10^{-21}$

## APPENDIX F: LINEAR ARRAY OF CONVERTERS

We discuss the different contributions to the optical noise starting from the expression for  $\bar{n}_f$ , Eq. (5) in the main text. When  $\Gamma_{\text{opt}} \ll \Gamma_{\nu}^*$ , the equation for the vibrational population splits into three different factors identified as thermal  $\bar{n}_{\text{th}}$ , dynamical backaction  $\bar{n}_{\text{dba}}$ , and quantum backaction noises  $\bar{n}_{\text{qba}}$ , respectively:

$$\bar{n}_f \simeq \bar{n}_{\text{th}} - \frac{\Gamma_{\text{opt}}}{\Gamma_{\nu}^*} \bar{n}_{\text{th}} + \frac{A^+}{\Gamma_{\nu}^*} \left( 1 - \frac{\Gamma_{\text{opt}}}{\Gamma_{\nu}^*} \right). \quad (\text{F1})$$

For sensing applications, it is enlightening to study how the contributions from the different noise terms are affected when considering an array of converters coherently illuminated by the IR field and the pump laser. We describe a linear array of  $N_{\text{conv}}$  optomechanical converters. For simplicity, we consider identical converters separated uniformly with a spacing  $d < \lambda_{\text{opt}} < \lambda_{\text{IR}}$  in order to avoid multiple maxima in the radiation pattern of the array. If all converters are excited in phase, the described configuration is known as the broadside configuration, and the maximum radiation is directed normal to the array axis. We assume that the optical pump power is split among the antennas, [56] so that the pump power per antenna is diluted according to  $|\alpha^{(i)}|^2 = (1/N_{\text{conv}})|\alpha^{(0)}|^2$  and the optomechanical coupling rate scales as  $1/N_{\text{conv}}$ , i.e.,  $A^{\pm(i)} = (1/N_{\text{conv}})A^{\pm(0)}$ . In this section, we use the superscripts (0) vs (i) to designate quantities computed under single-converter operation vs array operation (both of them refer to single-converter quantities).

### 1. Thermal regime

If the backaction effects are weak at a single-converter level  $\Gamma_{\text{opt}}^{(0)} < \Gamma_{\nu}^{*(0)}$ , the power dilution leads to  $\Gamma_{\text{opt}}^{(i)} \ll \Gamma_{\nu}^{*(i)}$ .



The expression of the final population  $\bar{n}_f$  [Eq. (F1)] shows that in this case thermal noise is the main contribution to the total noise.

In the far field, constructive interference among the fields emitted from individual antennas sharpens the pattern of coherent radiation [44] so that the total IR converted signal in this direction scales as [57]  $S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(N_{\text{conv}})} = (\text{array factor})^2 \cdot S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(i)}$ , which results in  $S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(N_{\text{conv}})} \simeq N_{\text{conv}}^2 S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(i)}$  along the direction of maximum radiation for a broadside array. On the contrary, if the converters are sufficiently spaced to avoid any near-field coupling, the thermal emission remains incoherent and quasi-isotropic.

We combine the factors related to the power dilution and to the directivity of the linear array to describe the SNR of the array in the regime dominated by thermal noise:

$$\text{SNR}^{(N_{\text{conv}})} \simeq N_{\text{conv}}^2 \cdot \frac{S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(i)}}{S_{\text{th}}^{\text{out},(i)}} = N_{\text{conv}}^2 \cdot \frac{1}{N_{\text{conv}}} \cdot \frac{S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(0)}}{S_{\text{th}}^{\text{out},(0)}} \quad (\text{F2})$$

## 2. Zero-temperature limit

In the case where the thermal population of the vibrational mode is negligible ( $\bar{n}_{\text{th}} \sim 0$ ), corresponding to back-action noise dominating over thermal noise, the incoming power dilution lowers equivalently the converted signal and the output noise per antenna, so that the SNR of a sufficiently large array in this regime is given by

$$\text{SNR}^{(N_{\text{conv}})} \simeq N_{\text{conv}}^2 \cdot \frac{S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(i)}}{S_{\text{ba}}^{\text{out},(i)}} = N_{\text{conv}}^2 \cdot 1 \cdot \frac{S_{\text{IR} \rightarrow \text{opt}}^{\text{out},(0)}}{S_{\text{ba}}^{\text{out},(0)}} \quad (\text{F3})$$

- [1] M. Tonouchi, *Cutting-Edge Terahertz Technology*, *Nat. Photonics* **1**, 97 (2007).
- [2] S. Ariyoshi, K. Nakajima, A. Saito, T. Taino, C. Otani, H. Yamada, S. Ohshima, J. Bae, and S. Tanaka, *Terahertz Response of NbN-Based Microwave Kinetic Inductance Detectors with Rewound Spiral Resonator*, *Supercond. Sci. Technol.* **29**, 035012 (2016).
- [3] J. Bueno, O. Yurduseven, S. J. C. Yates, N. Llombart, V. Murugesan, D. J. Thoen, A. M. Baryshev, A. Neto, and J. J. A. Baselmans, *Full Characterisation of a Background Limited Antenna Coupled KID over an Octave of Bandwidth for THz Radiation*, *Appl. Phys. Lett.* **110**, 233503 (2017).
- [4] F. Sizov, *Terahertz Radiation Detectors: The State-of-the-Art*, *Semicond. Sci. Technol.* **33**, 123001 (2018).
- [5] A. Rogalski, M. Kopytko, and P. Martyniuk, *Two-Dimensional Infrared and Terahertz Detectors: Outlook and Status*, *Appl. Phys. Rev.* **6**, 021316 (2019).
- [6] P. Roelli, *Christophe Galland, Nicolas Piro, and Tobias J. Kippenberg, Molecular Cavity Optomechanics as a Theory of Plasmon-Enhanced Raman Scattering*, *Nat. Nanotechnol.* **11**, 164 (2016).
- [7] G. N. Gol'tsman, O. Okunev, G. Chulkova, A. Lipatov, A. Semenov, K. Smirnov, B. Voronov, A. Dzardanov, C. Williams, and R. Sobolewski, *Picosecond Superconducting Single-Photon Optical Detector*, *Appl. Phys. Lett.* **79**, 705 (2001).
- [8] R. H. Hadfield, *Single-Photon Detectors for Optical Quantum Information Applications*, *Nat. Photonics* **3**, 696 (2009).
- [9] L. Tian and H. Wang, *Optical Wavelength Conversion of Quantum States with Optomechanics*, *Phys. Rev. A* **82**, 053806 (2010).
- [10] C. Dong, V. Fiore, M. C. Kuzyk, and H. Wang, *Optomechanical Dark Mode*, *Science* **338**, 1609 (2012).
- [11] J. T. Hill, A. H. Safavi-Naeini, J. Chan, and O. Painter, *Coherent Optical Wavelength Conversion via Cavity Optomechanics*, *Nat. Commun.* **3**, 1196 (2012).
- [12] T. A. Palomaki, J. W. Harlow, J. D. Teufel, R. W. Simmonds, and K. W. Lehnert, *Coherent State Transfer between Itinerant Microwave Fields and a Mechanical Oscillator*, *Nature (London)* **495**, 210 (2013).
- [13] J. Bochmann, A. Vainsencher, D. D. Awschalom, and A. N. Cleland, *Nanomechanical Coupling between Microwave and Optical Photons*, *Nat. Phys.* **9**, 712 (2013).
- [14] R. W. Andrews, R. W. Peterson, T. P. Purdy, K. Cicak, R. W. Simmonds, C. A. Regal, and K. W. Lehnert, *Bidirectional and Efficient Conversion between Microwave and Optical Light*, *Nat. Phys.* **10**, 321 (2014).
- [15] M. Forsch, R. Stockill, A. Wallucks, I. Marinkovic, C. Gärtner, R. A. Norte, F. van Otten, A. Fiore, K. Srinivasan, and S. Gröblacher, *Microwave-to-Optics Conversion Using a Mechanical Oscillator in Its Quantum Ground State*, *Nat. Phys.* **16**, 69 (2020).
- [16] C. Belacel, Y. Todorov, S. Barbieri, D. Gacemi, I. Favero, and C. Sirtori, *Optomechanical Terahertz Detection with Single Meta-Atom Resonator*, *Nat. Commun.* **8**, 1578 (2017).
- [17] R. W. Boyd and D. Prato, *Nonlinear Optics* (Elsevier Science, New York, 2008).
- [18] K. Karstad, A. Stefanov, M. Wegmuller, H. Zbinden, N. Gisin, T. Aellen, M. Beck, and J. Faist, *Detection of Mid-IR Radiation by Sum Frequency Generation for Free Space Optical Communication*, *Opt. Lasers Eng. Opt. Switzerland*, **43**, 537 (2005).
- [19] P. Tidemand-Lichtenberg, J. S. Dam, H. V. Andersen, L. Høgstedt, and C. Pedersen, *Mid-Infrared Upconversion Spectroscopy*, *J. Opt. Soc. Am. B* **33**, D28 (2016).
- [20] Yu.-P. Tseng, C. Pedersen, and P. Tidemand-Lichtenberg, *Upconversion Detection of Long-Wave Infrared Radiation from a Quantum Cascade Laser*, *Opt. Mater. Express* **8**, 1313 (2018).
- [21] G. Herzberg, *Molecular Spectra and Molecular Structure II. Infrared and Raman Spectra of Polyatomic Molecules* (Van Nostrand, Princeton, 1960).
- [22] N. B. Colthup, L. H. Daly, and S. E. Wiberley, *Introduction to Infrared and Raman Spectroscopy*: 3rd ed. (Elsevier Science, New York, 1990).
- [23] A. Shalabney, J. George, J. Hutchison, G. Pupillo, C. Genet, and T. W. Ebbesen, *Coherent Coupling of Molecular Resonators with a Microcavity Mode*, *Nat. Commun.* **6**, 5981 (2015).

- [24] C. Cohen-Tannoudji, J. Dupont-Roc, and G. Grynberg, *Atom-Photon Interactions: Basic Processes and Applications* (Wiley, New York, 1992).
- [25] B. Metzger, E. Muller, J. Nishida, B. Pollard, M. Hentschel, and M. B. Raschke, *Purcell-Enhanced Spontaneous Emission of Molecular Vibrations*, *Phys. Rev. Lett.* **123**, 153001 (2019).
- [26] I. Wilson-Rae, N. Nooshi, J. Dobrindt, T. J. Kippenberg, and W. Zwerger, *Cavity-Assisted Backaction Cooling of Mechanical Resonators*, *New J. Phys.* **10**, 095007 (2008).
- [27] M. O. Scully and M. S. Zubairy, *Quantum Optics* (Cambridge University Press, Cambridge, England, 1997).
- [28] A. Schliesser and T. J. Kippenberg, *Cavity Optomechanics with Whispering-Gallery Mode Optical Micro-resonators*, in *Advances in Atomic, Molecular, and Optical Physics* (Elsevier, New York, 2010), Vol. 58, pp. 207–323.
- [29] M. K. Schmidt, R. Esteban, A. Gonzalez-Tudela, G. Giedke, and J. Aizpurua, *Quantum Mechanical Description of Raman Scattering from Molecules in Plasmonic Cavities*, *ACS Nano* **10**, 6291 (2016).
- [30] M. K. Schmidt, R. Esteban, F. Benz, J. J. Baumberg, and J. Aizpurua, *Linking Classical and Molecular Optomechanics Descriptions of SERS*, *Faraday Discuss.* **205**, 31 (2017).
- [31] In this expression, we did not include the frequency dependence of the photonic density of state in free space nor the factors related to the specific optical design used to couple light in and out of the structure.
- [32] E. B. Wilson, J. C. Decius, and P. C. Cross, *Molecular Vibrations: The Theory of Infrared and Raman Vibrational Spectra* (Courier, New York, 1980).
- [33] S. Haroche and J.-M. Raimond, *Exploring the Quantum: Atoms, Cavities, and Photons* (Oxford University Press, New York, 1997).
- [34] W. Zhu and K. B. Crozier, *Quantum Mechanical Limit to Plasmonic Enhancement as Observed by Surface-Enhanced Raman Scattering*, *Nat. Commun.* **5**, 5228 (2014).
- [35] F. Benz, M. K. Schmidt, A. Dreismann, R. Chikkaraddy, Y. Zhang, A. Demetriadou, C. Carnegie, H. Ohadi, B. de Nijs, R. Esteban, J. Aizpurua, and J. J. Baumberg, *Single-Molecule Optomechanics in “Picocavities,”* *Science* **354**, 726 (2016).
- [36] L. V. Brown, X. Yang, K. Zhao, B. Y. Zheng, P. Nordlander, and N. J. Halas, *Fan-Shaped Gold Nanoantennas above Reflective Substrates for Surface-Enhanced Infrared Absorption (SEIRA)*, *Nano Lett.* **15**, 1272 (2015).
- [37] F. Neubrech, C. Huck, K. Weber, A. Pucci, and H. Giessen, *Surface-Enhanced Infrared Spectroscopy Using Resonant Nanoantennas*, *Chem. Rev.* **117**, 5110 (2017).
- [38] F. Schreiber, *Structure and Growth of Self-Assembling Monolayers*, *Prog. Surf. Sci.* **65**, 151 (2000).
- [39] J. C. Love, L. A. Estroff, J. K. Kriebel, R. G. Nuzzo, and G. M. Whitesides, *Self-Assembled Monolayers of Thiolates on Metals as a Form of Nanotechnology*, *Chem. Rev.* **105**, 1103 (2005).
- [40] M. Long, A. Gao, P. Wang, H. Xia, C. Ott, C. Pan, Y. Fu, E. Liu, X. Chen, W. Lu, T. Nilges, J. Xu, X. Wang, W. Hu, and F. Miao, *Room Temperature High-Detectivity Mid-infrared Photodetectors Based on Black Arsenic Phosphorus*, *Sci. Adv.* **3**, e1700589 (2017).
- [41] X. Chen, X. Lu, B. Deng, O. Sinai, Y. Shao, C. Li, S. Yuan, V. Tran, K. Watanabe, T. Taniguchi, D. Naveh, L. Yang, and F. Xia, *Widely Tunable Black Phosphorus Mid-infrared Photodetector*, *Nat. Commun.* **8**, 1 (2017).
- [42] H. M. Doleman, E. Verhagen, and A. F. Koenderink, *Antenna–Cavity Hybrids: Matching Polar Opposites for Purcell Enhancements at Any Linewidth*, *ACS Photonics* **3**, 1943 (2016).
- [43] B. Gurlek, V. Sandoghdar, and D. Martin-Cano, *Manipulation of Quenching in Nanoantenna–Emitter Systems Enabled by External Detuned Cavities: A Path to Enhance Strong-Coupling*, *ACS Photonics* **5**, 456 (2018).
- [44] D. Dregely, R. Taubert, J. Dorfmueller, R. Vogelgesang, K. Kern, and H. Giessen, *3D Optical Yagi–Uda Nanoantenna Array*, *Nat. Commun.* **2**, 267 (2011).
- [45] S. Busschaert, N. Flöry, S. Papadopoulos, M. Parzefall, S. Heeg, and L. Novotny, *Beam Steering with a Nonlinear Optical Phased Array Antenna*, *Nano Lett.* **19**, 6097 (2019).
- [46] B. Schwarz, P. Reininger, D. Ristanic, H. Detz, A. M. Andrews, W. Schrenk, and G. Strasser, *Monolithically Integrated Mid-infrared Lab-on-a-Chip Using Plasmonics and Quantum Cascade Structures*, *Nat. Commun.* **5**, 4085 (2014).
- [47] H. Lin, Z. Luo, T. Gu, L. C. Kimerling, K. Wada, A. Agarwal, and J. Hu, *Mid-infrared Integrated Photonics on Silicon: A Perspective*, *Nanophotonics* **7**, 393 (2018).
- [48] A. Tittl, A. Leitis, M. Liu, F. Yesilkoy, D.-Y. Choi, D. N. Neshev, Y. S. Kivshar, and H. Altug, *Imaging-Based Molecular Barcoding with Pixelated Dielectric Metasurfaces*, *Science* **360**, 1105 (2018).
- [49] F. Yesilkoy, E. R. Arvelo, Y. Jahani, M. Liu, A. Tittl, V. Cevher, Y. Kivshar, and H. Altug, *Ultrasensitive Hyper-spectral Imaging and Biodetection Enabled by Dielectric Metasurfaces*, *Nat. Photonics* **13**, 390 (2019).
- [50] G. Albrecht, S. Kaiser, H. Giessen, and M. Hentschel, *Refractory Plasmonics without Refractory Materials*, *Nano Lett.* **17**, 6402 (2017).
- [51] Zenodo data package for the manuscript, <https://doi.org/10.5281/zenodo.3709007>.
- [52] C. G. Tetsassi Feugmo and V. Liégeois, *Analyzing the Vibrational Signatures of Thiophenol Adsorbed on Small Gold Clusters by DFT Calculations*, *Chem. Phys. Chem.* **14**, 1633 (2013).
- [53] E. Le Ru and P. Etchegoin, *Principles of Surface-Enhanced Raman Spectroscopy: and Related Plasmonic Effects* (Elsevier, New York, 2008).
- [54] *CRC Handbook of Chemistry and Physics*, 87th ed., edited by D. R. Lide (CRC Press, Boca Raton, FL, 2006).
- [55] A. Sundaramurthy, K. B. Crozier, G. S. Kino, D. P. Fromm, P. J. Schuck, and W. E. Moerner, *Field Enhancement and Gap-Dependent Resonance in a System of Two Opposing Tip-to-Tip Au Nanotriangles*, *Phys. Rev. B* **72**, 165409 (2005).
- [56] The diffraction limit for both beams being largely different, we note that multiple converters fit under a focused IR spot. In that case, the IR power per converter would not scale down.
- [57] C. A. Balanis, *Antenna Theory: Analysis and Design* (Wiley-Interscience, New York, 2005).