Coupling theory of quasinormal modes for lossy and dispersive plasmonic nanoresonators

Can Tao,[1] Junda Zhu ⬤,[2] Ying Zhong,[3] and Haitao Liu ⬤[1,*]

[1]*Tianjin Key Laboratory of Micro-scale Optical Information Science and Technology, Institute of Modern Optics,*
College of Electronic Information and Optical Engineering, Nankai University, Tianjin 300350, China
[2]*The MOE Key Laboratory of Weak Light Nonlinear Photonics, TEDA Institute of Applied Physics and School of Physics,*
Nankai University, Tianjin 300457, China
[3]*State Key Laboratory of Precision Measurement Technology and Instruments, Tianjin University, Tianjin 300072, China*

A quasinormal mode (QNM) coupling theory is proposed for a system of coupled resonators. The theory is generally applicable to systems with energy loss and arbitrary frequency dispersion of permittivity such as plasmonic nanoresonators. The QNM coupling coefficients in the theory are obtained with first-principle calculations without fitting the full-wave numerical results or experimental data. The theory is built up based on the Green's-function formalism, which shows that the coupling between resonators originates from the mutual excitation of scattered field of each resonator. With the theory, the source-excited scattered field or the source-free QNMs of the coupled system can be obtained as an expansion in terms of QNMs of each individual resonator in the system. The theory provides an intuitive physical picture and also shows superior computational efficiency compared with the full-wave numerical method if the scattered field of each resonator in the system can be expanded with a small set of QNMs. The theory is tested against full-wave numerical method and is used to design metallic nanoantennas with a broadband enhancement of the spontaneous emission rate.

I. INTRODUCTION

Owing to the support of localized surface plasmon resonance (LSPR), metallic nanoresonators (MNRs) can confine the electromagnetic field within a deep subwavelength region far beyond the diffraction limit [1]. Based on the property, MNRs have been intensively studied and widely applied in various fields such as spontaneous emission enhancement [2], single-molecule sensing [3,4], photodetector [5], heat transfer enhancement [6], super-resolution imaging [7], and nanolaser [8,9]. With coupled MNRs, some interesting physical properties are found. For example, the nanogap between two MNRs can support the gap surface plasmon polariton, which allows energy to be concentrated in the nanogap and thus further enhances the electromagnetic field and the spontaneous emission rate [10]. Broadband enhancement of electromagnetic field [11,12] can be achieved in coupled MNRs due to the support of multiple resonant modes. Directional far-field radiation can be realized with coupled MNRs [13] based on the constructive interference of the scattered fields of each MNR. Thanks to these properties, coupled MNRs possess superior performance than a single MNR, in terms of enhanced Raman-scattering spectroscopy [14], single-molecule fluorescence enhancement [15,16], nonlinear response enhancement [17,18], etc. Different from single MNR, the resonance frequencies and spectral response shape of coupled MNRs depend on not only the shape, size, and material of MNRs, but also the number and spacing of MNRs. Based on the dependence of the scattering spectrum of the coupled

MNRs on the MNRs' spacing, a plasmon ruler is proposed to measure the length of labeled biomolecules [19,20]. With the dependence of the scattering spectrum of the coupled MNRs on the number of MNRs, plasmonic imaging is achieved to examine the spatial distribution of labeled biomolecules [21,22]. In addition, a metamaterial composed of coupled MNRs can realize plasmon-induced transparency [23,24] based on the coupling between bright and dark modes. Due to the existence of absorption and radiation loss, coupled MNRs are typical non-Hermitian systems and thus can support an exceptional point [25–27].

By considering the coupling between resonant modes or waveguide modes of a system, mode coupling theory can provide an intuitive physical picture to effectively guide the design of devices and can reduce the computational amount [28–31]. Therefore, it has attracted great attention of researchers. For lossless coupled MNRs consisting of spherical MNRs, the dependence of the resonance frequency of the system on the coupling between the plasmon modes of each MNR is analyzed based on a plasmon hybridization model [32–34]. For coupled MNRs consisting of arbitrarily shaped MNRs, the dependence of the scattered field of coupled MNRs on the coupling between MNRs is analyzed by building up a coupling theory based on the boundary element method under the electrostatic approximation [35]. For lossless coupled MNRs whose permittivity is described by a single-pole Lorentz model, a coupled-mode theory [29,36,37] is established by considering the coupling between a single resonant mode of each MNR under a generalized tight-binding approximation. By treating the loss and permittivity dispersion as perturbations, the classical time-domain coupled mode theory (TCMT) can be derived from a consideration of

*Corresponding author: liuht@nankai.edu.cn

045430-1

the power distribution in the coupled system [28,38,39]. The classical TCMT is sometimes used as a phenomenological theory, where the model parameters are obtained by fitting full-wave numerical results or experimental data [27,40,41].

For an open system with absorption or scattering loss, its supported resonant mode is a quasinormal mode (QNM) with a complex resonance/eigen frequency [42–47]. For a lossless closed system, QNM becomes normal mode with a real resonance frequency [46,47]. The LSPR supported by MNRs can be rigorously defined as the QNM [48]. Based on the QNM expansion formalism, TCMT for the coupling between QNMs of a single photonic crystal resonator and waveguide modes of photonic crystal waveguides is built up, where the material permittivity is nondispersive [49]. For a dispersive and lossy system composed of a single resonant cavity and multiple ports, a QNM expansion of the scattering matrix is obtained by building up a coupled-mode theory [31]. Based on a direct expansion-projection procedure upon the basis of QNMs of single resonators, a coupling theory for the QNMs of two nondispersive coupled resonators is established [50]. Recently, the coupling theories for the QNMs of lossy and dispersive coupled resonators immersed in a homogeneous background medium are proposed by using the QNM expansion formalism and a multiple scattering picture [51,52]. For coupled MNRs, the finite loss of metal needs to be considered [53,54], and the permittivity of metal can be modeled with functions that are more complex than the single-pole Lorentz model, such as the multiple-pole Lorentz model [55] and the fractional function with multipoles [56,57]. For coupled MNRs with finite energy loss and arbitrary frequency dispersion of permittivity, a rigorous coupled-mode theory of QNMs based on the first principles of Maxwell's equations has not been established yet.

In addition to the coupling theories of QNMs (with an eigenfrequency), the coupling theories of the generalized normal modes (with an eigenpermittivity) have been established for lossy and dispersive coupled systems [58–60] with a homogeneous distribution of permittivity inside each isolated scatterer and in the background. The expansion of the source-excited field in terms of the generalized normal modes is analytic with respect to the scatterer permittivity [61]. In contrast, the expansion of the source-excited field in terms of the QNMs is analytic with respect to the excitation frequency [47,62–66]. This frequency analyticity is important for many applications, such as for calculating the spectral and the temporal response of the system [65], for calculating the nonlinear response [67], and for the quantization of electromagnetic field in lossy and dispersive resonators [68].

In this paper, a general and rigorous QNM coupling theory based on the first principles of Maxwell's equations is proposed for coupled MNRs. Compared with the existing mode coupling theory, the main advantage of our theory is that it is generally applicable to coupled-MNR systems with loss and arbitrary frequency dispersion of permittivity, and the QNM coupling coefficients in the theory are calculated based on the first principles, without fitting the full wave numerical results or experimental data.

First, for a general coupled-MNR system, we analyze the coupling mechanism based on the Green's-function formalism and find that the scattered field of the coupled MNRs can be obtained as the superposition of the scattered fields of each individual MNR in the system. For the scattered field of each individual MNR, the corresponding incident field consists of two parts, one part is the background field excited by an external source, and the other part is the scattered field from the other MNRs. Based on this physical picture and by expressing the scattered field of each individual MNR with the QNM expansion formalism [47,62,63], we then build up a QNM coupling theory to describe the mutual coupling between the QNMs of each MNR in the coupled-MNR system. With the proposed QNM coupling theory, the scattered field of the coupled MNRs can be solved analytically when the external source and the QNMs of each MNR are known. Therefore, compared with the full-wave numerical method, the proposed theory is physically intuitive and also has higher computational efficiency for repeated calculations when changing the source frequency, the MNR number, and the relative position between MNRs. Furthermore, by removing the external source, the QNMs supported by the coupled MNRs can be obtained as a superposition of the QNMs supported by each MNR. The validity of the theory is tested against full-wave numerical results. With the theory, several coupled-MNR structures are designed to achieve a broadband enhancement of the spontaneous emission rate.

This paper is organized as follows. In Sec. II, the existing QNM expansion formalism is briefly reviewed. In Sec. III, the QNM coupling theory is introduced. In Sec. IV, the theory is tested against full-wave numerical results and is used to design coupled-MNR devices. Conclusions are summarized in Sec. V.

II. A BRIEF REVIEW OF THE QNM-EXPANSION FORMALISM

As a preparation for deriving the QNM coupling theory in the next section, here we provide a brief review of the QNM expansion formalism [47,62–66]. The QNMs are the eigensolutions of the source-free Maxwell's equations,

$$\nabla \times \tilde{\mathbf{E}}_m = i\tilde{\omega}_m \boldsymbol{\mu}(\mathbf{r}, \tilde{\omega}_m) \cdot \tilde{\mathbf{H}}_m,$$
$$\nabla \times \tilde{\mathbf{H}}_m = -i\tilde{\omega}_m \boldsymbol{\varepsilon}(\mathbf{r}, \tilde{\omega}_m) \cdot \tilde{\mathbf{E}}_m \qquad (1)$$

and satisfy the outgoing-wave condition at infinity. In Eq. (1), $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$ represent the permittivity and permeability tensors, respectively, and the temporal dependence $\exp(-i\tilde{\omega}_m t)$ of the field is assumed. The international system of units (SI) is adopted for the formula in this paper. Due to the energy loss of the system, the QNM lifetime τ_m takes a finite value. Therefore, the resonance frequency $\tilde{\omega}_m$ takes a complex value, and its imaginary part $\text{Im}(\tilde{\omega}_m) = -1/(2\tau_m)$ represents the energy loss. Then the imaginary part of the wave number $k = \tilde{\omega}_m/c$ (c being the speed of light in vacuum) in free space takes a negative value because of $\text{Im}(\tilde{\omega}_m) < 0$. Thus, the QNM field $\tilde{\boldsymbol{\Psi}}_m(\mathbf{r}) = [\tilde{\mathbf{E}}_m, \tilde{\mathbf{H}}_m]$ diverges in the form of $\exp(ikr)/r$ (r being the distance to the scatterer) at infinity in the free space [47]. The quality factor of the QNM is defined as [47] $Q_m = \text{Re}(\tilde{\omega}_m)/[-2\text{Im}(\tilde{\omega}_m)]$, which represents the number of cycles of the light field oscillation within the mode lifetime τ_m.

In the presence of an excitation source, the electromagnetic field $\Psi(\mathbf{r}, \omega) = [\mathbf{E}, \mathbf{H}]$ of the system satisfies the frequency-domain Maxwell's equations,

$$\nabla \times \mathbf{E} = i\omega\boldsymbol{\mu}(\mathbf{r}, \omega) \cdot \mathbf{H},$$
$$\nabla \times \mathbf{H} = -i\omega\boldsymbol{\varepsilon}(\mathbf{r}, \omega) \cdot \mathbf{E} + \mathbf{J}(\mathbf{r}, \omega), \qquad (2)$$

where $\mathbf{J}$ represents the electric current source with a real angular frequency ω. By assuming that the QNMs of the system form a complete set of basis [45,57,63,65], the $\Psi(\mathbf{r}, \omega)$ can then be expressed as a linear combination of a series of QNMs,

$$\Psi(\mathbf{r}, \omega) \approx \sum_{m=1}^{M} \alpha_m(\omega)\tilde{\Psi}_m(\mathbf{r}), \qquad (3)$$

where the approximate equality becomes equality for $M \to \infty$. According to the QNM-expansion formalism and for reciprocal medium ($\boldsymbol{\varepsilon}^{\mathrm{T}} = \boldsymbol{\varepsilon}, \boldsymbol{\mu}^{\mathrm{T}} = \boldsymbol{\mu}$, with the superscript denoting the tensor transposition), the expansion coefficient $\alpha_m(\omega)$ can be expressed as [62–64,66]

$$\alpha_m(\omega) = \frac{\iiint_{R^3} \mathbf{J}(\mathbf{r}, \omega) \cdot \tilde{\mathbf{E}}_m(\mathbf{r})d^3\mathbf{r}}{i(\omega - \tilde{\omega}_m)F_m}, \qquad (4)$$

with

$$F_m = \iiint_{R^3}$$
$$\times \left\{ \tilde{\mathbf{E}}_m \cdot \frac{\partial[\omega\boldsymbol{\varepsilon}(\omega)]}{\partial\omega} \cdot \tilde{\mathbf{E}}_m - \tilde{\mathbf{H}}_m \cdot \frac{\partial[\omega\boldsymbol{\mu}(\omega)]}{\partial\omega} \cdot \tilde{\mathbf{H}}_m \right\}_{\omega=\tilde{\omega}_m} d^3\mathbf{r}, \qquad (5)$$

the integral being performed over the whole space R^3. If the quality factor Q_m of the mth-order QNM is of finite value, F_m then takes a complex value, and its numerical calculation is explained in Supplemental Material Sec. S1 [69]. For $Q_m \to \infty$, which requires the medium to be lossless, one can obtain $\tilde{\mathbf{E}}_m = \tilde{\mathbf{E}}_m^*$ and $\tilde{\mathbf{H}}_m = -\tilde{\mathbf{H}}_m^*$. Then the QNM becomes normal mode, and F_m takes a real value and can represent the electromagnetic energy of the normal mode [47,70]. Therefore, if Q_m takes a large value, $|F_m|$ can be used to represent the energy of the QNM in an approximate manner. Thus F_m is called pseudoenergy in this paper.

Note that Eq. (4) is generally applicable to a system with energy loss and arbitrary frequency dispersion of $\boldsymbol{\varepsilon}$ and $\boldsymbol{\mu}$. For more details, for a system whose permittivity dispersion can be modeled with the quite general N-pole Lorentz model, the orthogonality of QNMs can be proved by introducing the auxiliary fields and using the reciprocity theorem of unconjugated form, and then the QNM expansion coefficient such as Eq. (4) can be obtained [65,66]. It should be noted that if different forms of auxiliary fields are used to express the excitation source $\mathbf{J}$, the corresponding expressions of QNM expansion coefficient will be different, but the convergence of the series ($M \to \infty$) given by Eq. (3) can be always guaranteed (although the convergence rate may be different) [66]. Equation (4) adopted in this paper is just one of these expressions of the QNM expansion coefficient. For a system with an arbitrary frequency dispersion of medium, Eq. (4) can

be derived from the Green's dyadic obtained via the Mittag-Leffler theorem [62] or the Keldyš theorem [63].

For the scattering problem, we assume that the total field excited by the current source satisfies Eq. (2), and consider the nonmagnetic medium ($\boldsymbol{\mu} = \mu_0$). The background field $\Psi^b = [\mathbf{E}^b, \mathbf{H}^b]$ excited by $\mathbf{J}(\mathbf{r}, \omega)$ is defined to satisfy $\nabla \times \mathbf{E}^b = i\omega\mu_0\mathbf{H}^b$, $\nabla \times \mathbf{H}^b = -i\omega\boldsymbol{\varepsilon}_b \cdot \mathbf{E}^b + \mathbf{J}$, where $\boldsymbol{\varepsilon}_b$ is the permittivity tensor of the background medium in the absence of any scatterer. Thus, the scattered field $\Psi^s = \Psi - \Psi^b$ satisfies the Maxwell's equations,

$$\nabla \times \mathbf{E}^s = i\omega\mu_0\mathbf{H}^s, \quad \nabla \times \mathbf{H}^s = -i\omega\boldsymbol{\varepsilon} \cdot \mathbf{E}^s - i\omega\Delta\boldsymbol{\varepsilon} \cdot \mathbf{E}^b, \qquad (6)$$

where $\Delta\boldsymbol{\varepsilon}(\mathbf{r}, \omega) = \boldsymbol{\varepsilon}(\mathbf{r}, \omega) - \boldsymbol{\varepsilon}_b(\mathbf{r}, \omega)$ is defined as the permittivity change of the scatterer. Equation (6) shows that the scattered field can be regarded as the field radiated by an equivalent current source $-i\omega\Delta\boldsymbol{\varepsilon}(\mathbf{r}, \omega) \cdot \mathbf{E}^b$ in the presence of the scatterer. Therefore, similar to Eq. (3), the scattered field Ψ^s can be expanded upon the basis of QNMs [Eq. (1)] of the system in the presence of the scatterer,

$$\Psi^s(\mathbf{r}, \omega) \approx \sum_{m=1}^{M} \alpha_m(\omega)\tilde{\Psi}_m(\mathbf{r}), \qquad (7)$$

with [66,71]

$$\alpha_m(\omega) = \frac{-\omega \iiint_{R^3} \tilde{\mathbf{E}}_m(\mathbf{r}) \cdot \Delta\boldsymbol{\varepsilon}(\mathbf{r}, \omega) \cdot \mathbf{E}^b(\mathbf{r})d^3\mathbf{r}}{(\omega - \tilde{\omega}_m)F_m}. \qquad (8)$$

In the following, we will provide a brief description on the numerical implementation of the QNM-expansion formalism. First, we need to numerically calculate the QNM electromagnetic fields $\tilde{\Psi}_m$ and eigenfrequencies $\tilde{\omega}_m$. For dispersive media, there are mainly two kinds of methods [47,72]. One is to transform the nonlinear eigenvalue problem of Eq. (1) into a linear eigenvalue problem or a polynomial eigenvalue problem by introducing the auxiliary fields [65,73,74]. Another method is the pole-search approach which is based on the fact that the resonance frequency $\tilde{\omega}_m$ is the complex frequency pole of the source-excited electromagnetic field [i.e., $\Psi(\mathbf{r}, \tilde{\omega}_m) \to \infty$ according to Eqs. (3) and (4)]. For example, $\tilde{\omega}_m$ can be solved with some iterative algorithms such as the linear interpolation iterative algorithm [75]. Or based on the residue theorem, all complex frequency poles within a contour can be found by calculating a contour integral and solving a set of nonlinear equations [76]. The $\tilde{\omega}_m$ can be also obtained by calculating the complex frequency poles of the scattering matrix [77,78]. In this paper, we adopt the linear interpolation iterative algorithm [75].

Since the resonance frequency takes a complex value, an analytical continuation of the relative permittivity in the complex frequency plane is required. For example, the frequency-dependent function of the relative permittivity can be obtained based on a microscopic description of the charge motion, which is called Lorentz model (its special case being the Drude model) [55]. In this paper, the Drude model of $\varepsilon_r(\omega) = \varepsilon_\infty - \omega_0^2/(\omega^2 + i\omega\gamma)$ is adopted for numerical calculation. The model parameters are obtained by fitting the experimental data [79] from visible to far-infrared band to
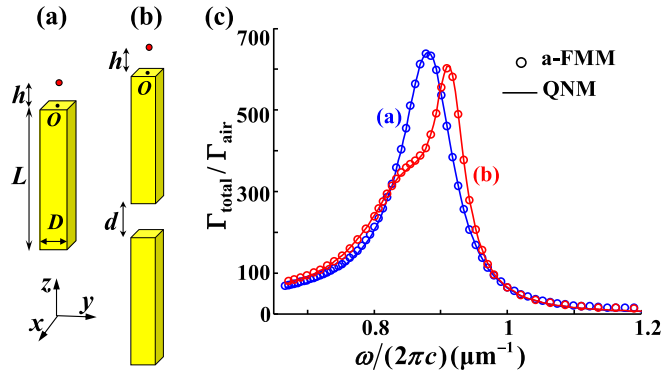
FIG. 1. (a) A gold single-wire nanoantenna (side length $D = 40$ nm for the rectangular cross section, length $L = 247$ nm) is excited by a z-polarized point current source (red point). The point source is away from the antenna end surface by $h = 10$ nm, and its projection on the antenna end surface is located at the center of the surface (this is assumed throughout the paper). (b) A dipole antenna consisting of two gold nanowires [with the same parameters as in (a)] is excited by a z-polarized point source (red point). The width of the nanogap between the two nanowires is $d = 100$ nm. (c) Spectrum of the total spontaneous emission rate Γ_{total} of the point source near the single-wire antenna (blue curves) and dipole antenna (red curves), where $\omega/(2\pi c) = 1/\lambda$ with λ being the wavelength. The circles and solid curves represent the full-wave a-FMM results and the QNM expansion formalism predictions, respectively.

be $\varepsilon_\infty = 8.842$, $2\pi c/\omega_0 = 0.164$ μm and $2\pi c/\gamma = 20.689$ μm. However, note that our theory to be proposed in next section is not limited to the Drude model but is generally applicable to an arbitrary frequency dispersion of permittivity.

With the QNMs obtained, then Eqs. (4) and (8) are used to calculate the QNM expansion coefficients of the radiation field and the scattered field, respectively. Finally, the radiation field and scattered field are obtained by using Eqs. (3) and (7), respectively, where the number M of QNMs takes a value that is large enough to meet the requirement of accuracy.

In the following, we will provide a numerical test of the afore-reviewed QNM expansion formalism. As shown in Fig. 1, we take the emission problem of metallic nanoantennas as an example. The MNR as sketched in Fig. 1(a) is a single gold nanowire, and all structural parameters are provided in the caption of the figure. Figure 1(b) shows a dipole antenna [5,10,80] whose two arms are just the nanowire shown in Fig. 1(a), and the spacing between the two arms is $d = 100$ nm. A z-polarized point electric current source is used to represent a fluorescent molecule or quantum dot. The point source can be expressed as $\mathbf{J}(\mathbf{r}) = \delta(\mathbf{r} - \mathbf{r}_0)\mathbf{z}$, with δ the Dirac delta function, $\mathbf{r}_0$ the position vector of the point source (as specified in Fig. 1), and $\mathbf{z}$ the unit vector along the z direction. The ratio of the spontaneous emission rate of the point source modified by the antenna to that of the point source in free space is equal to the ratio of the corresponding radiated powers [81]. Therefore, in the following text the spontaneous emission rate just refers to the radiated power and they are no longer distinguished for the sake of simplicity. The total spontaneous emission rate of the point source can be expressed as $\Gamma_{total} = -\text{Re}[E_z(\mathbf{r}_0)]/2$, where $\text{Re}[E_z(\mathbf{r}_0)]$ means the real part of the z component of the electric-field
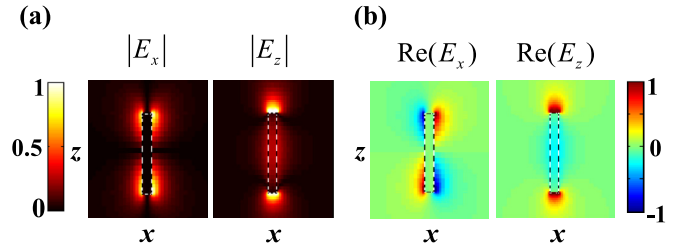

FIG. 2. Electric-field distribution [in the plane $y = 0$, the coordinate system being shown in Fig. 1(a)] of the dipolelike QNM [with a complex resonance frequency $\tilde{\omega}_0/(2\pi c) = 0.8842 - 0.0477i$ μm^{-1}] supported by the single-wire antenna shown in Fig. 1(a). Unless otherwise specified, the QNM fields shown in this paper always satisfy the normalization of $E_z(0, 0, 0) = 1$. (a) Distribution of the modulus of the x component $|E_x|$ (left) and z component $|E_z|$ (right) of the electric field. (b) Distribution of the real part of the x component $\text{Re}(E_x)$ (left) and z component $\text{Re}(E_z)$ (right) of the electric field. The superimposed dotted lines in the figure represent the antenna boundaries.

vector at the source position. To characterize the enhancement of the spontaneous emission rate, a Purcell factor is defined as $\Gamma_{total}/\Gamma_{air}$ [82,83], where $\Gamma_{air} = \eta_0 k_0^2 n_a/(12\pi)$ is the spontaneous emission rate of the point source in air, $n_a = 1$ being the air refractive index, η_0 and $k_0 = 2\pi/\lambda$ being the wave impedance and wave number in vacuum, respectively.

As shown in Fig. 1(c), the spectrum of $\Gamma_{total}/\Gamma_{air}$ is calculated with the QNM expansion formalism of Eqs. (3) and (4) (solid curves) and a full-wave aperiodic Fourier modal method (a-FMM) [84,85] (circles), respectively. For the single-wire antenna, we only consider one QNM with the highest Q factor [with a complex resonance frequency $\tilde{\omega}_0/(2\pi c) = 0.8842 - 0.0477i$ μm^{-1}] within the spectrum range of interest to expand the electromagnetic field excited by the point source. The field distribution of this QNM is shown in Fig. 2. For the dipole antenna, we consider two QNMs with the highest Q factors [with complex resonance frequencies $\tilde{\omega}_1/(2\pi c) = 0.8495 - 0.0697i$ μm^{-1}, $\tilde{\omega}_2/(2\pi c) = 0.9142 - 0.0274i$ μm^{-1}] to expand the electromagnetic field. Figure 1(c) shows that the predictions by the QNM expansion formalism are consistent well with the a-FMM calculation results, which confirms the validity and the high accuracy of the former. It can be seen that, for the dipole antenna composed of two coupled nanowires, there are two resonance peaks in the spectrum of Γ_{total}, whose corresponding frequencies are the real parts of $\tilde{\omega}_1$ and $\tilde{\omega}_2$. However, for the single-wire antenna, there is only one resonance peak in the spectrum, whose frequency is the real part of $\tilde{\omega}_0$.

For arbitrary resonant structures with material dispersion and loss, the QNM expansion formalism can enable an analysis of the frequency response of the system from the perspective of eigenmodes with an intuitive physical picture. In addition, once the QNMs of the structure are solved, the QNM expansion formalism can provide the source-excited field that is analytical with respect to the frequency, polarization and spatial distribution of the excitation source [see Eq. (4)]. Therefore, it is not necessary to repeatedly solve the Maxwell's equations when scanning the parameters of the excitation source, so that the computational amount can be greatly reduced [for the calculation in Fig. 1(c), for instance].

However, in the QNM expansion formalism, the determination of the QNMs of the system still requires a numerical solution of Maxwell's equations. Therefore, when the coupled-MNRs' structural parameters (such as the number and relative positions of MNRs) are changed, a repetition of solving Maxwell's equations is needed to obtain QNMs of the coupled MNRs, which will require a large amount of computation. On the other hand, determining these structural parameters is usually the key for the design of coupled-MNR devices [17,20,23,25,80]. Besides, a numerical solution of the QNMs of the coupled MNRs cannot clarify the mechanism of the coupling between the MNRs, and thus is not conducive to guiding the design of devices. In response to these problems, in the following sections we will propose a general QNM coupling theory for coupled-MNR systems with loss and arbitrary frequency dispersion of the permittivity.

III. QNM COUPLING THEORY FOR A SYSTEM OF COUPLED METALLIC NANORESONATORS

A. QNM coupling theory for the scattered field of the coupled-MNR system

We first analyze the coupling mechanism in coupled-MNRs based on the Green's-function formalism. For this purpose, Eq. (6) can be rewritten as

$$\nabla \times \mathbf{E}^s = i\omega\mu_0\mathbf{H}^s,$$
$$\nabla \times \mathbf{H}^s = -i\omega\boldsymbol{\varepsilon}_b \cdot \mathbf{E}^s - i\omega\Delta\boldsymbol{\varepsilon} \cdot (\mathbf{E}^s + \mathbf{E}^b), \quad (9)$$

where the nonmagnetic medium ($\mu = \mu_0$) is considered. Equation (9) indicates that the scattered field of the system can be also regarded as the field radiated by an equivalent current source $-i\omega\Delta\boldsymbol{\varepsilon} \cdot (\mathbf{E}^s + \mathbf{E}^b)$ in the absence of the scatterers of MNRs [which is different from Eq. (6)]. The Green's-function tensor $\mathbf{G}(\mathbf{r}, \mathbf{r}'; \omega)$ in the absence of the scatterers of MNRs is defined to satisfy

$$\nabla \times \nabla \times \mathbf{G} - \omega^2\mu_0\boldsymbol{\varepsilon}_b \cdot \mathbf{G} = i\omega\mu_0\mathbf{I}\delta(\mathbf{r} - \mathbf{r}'), \quad (10)$$

with $\mathbf{I}$ denoting the unit tensor. Thus, the electric-field vector of the scattered field can be expressed as [86]

$$\mathbf{E}^s(\mathbf{r}, \omega) = -i\omega \iiint_{R^3} \mathbf{G}(\mathbf{r}, \mathbf{r}'; \omega) \cdot \Delta\boldsymbol{\varepsilon}(\mathbf{r}', \omega)$$
$$\cdot [\mathbf{E}^s(\mathbf{r}', \omega) + \mathbf{E}^b(\mathbf{r}', \omega)]d^3\mathbf{r}'. \quad (11)$$

Equation (11) forms an integral equation with the $\mathbf{E}^s$ as the unknown. For an open system, $\mathbf{E}^s$ satisfies the outgoing-wave condition at infinity, and can be uniquely determined with Eq. (11).

As shown in Fig. 3, a coupled-MNR system composed of P arbitrary MNRs is considered, where $\boldsymbol{\varepsilon}_p$ represents the permittivity distribution in the presence of the pth MNR, and $\boldsymbol{\varepsilon}_b$ represents the permittivity distribution of the background medium in the absence of the scatterers of the MNRs. The scatterer permittivity change can be expressed as $\Delta\boldsymbol{\varepsilon} = \sum_{p=1}^{P} \Delta\boldsymbol{\varepsilon}_p$, where $\Delta\boldsymbol{\varepsilon}_p = \boldsymbol{\varepsilon}_p - \boldsymbol{\varepsilon}_b$ ($\Delta\boldsymbol{\varepsilon}_p = 0$ outside the pth MNR) is the permittivity change of the pth MNR. By substi-
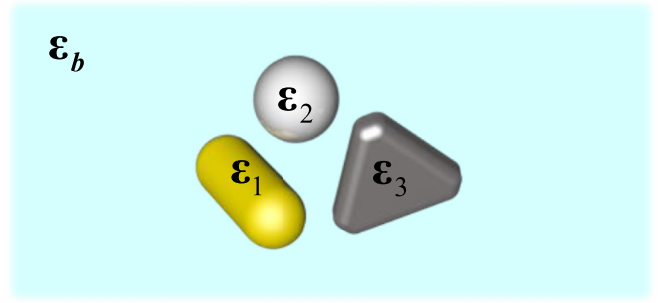

FIG. 3. Schematic diagram of the coupled-MNR system studied in this paper. The system is composed of P MNRs of arbitrary shape and material ($P = 3$ in the figure). The permittivity distribution of the background medium is denoted by $\boldsymbol{\varepsilon}_b$, and that of the pth MNR is denoted by $\boldsymbol{\varepsilon}_p(p = 1, 2, \ldots, P)$.

tuting the expression of $\Delta\boldsymbol{\varepsilon}$ into Eq. (11), $\mathbf{E}^s$ can be expressed as

$$\mathbf{E}^s(\mathbf{r}, \omega) = \sum_{p=1}^{P} \mathbf{E}_p^s(\mathbf{r}, \omega), \quad (12)$$

with

$$\mathbf{E}_p^s(\mathbf{r}, \omega) = -i\omega \iiint_{R^3} \mathbf{G}(\mathbf{r}, \mathbf{r}'; \omega) \cdot \Delta\boldsymbol{\varepsilon}_p(\mathbf{r}', \omega)$$
$$\cdot \left[\mathbf{E}_p^s(\mathbf{r}', \omega) + \sum_{q \neq p} \mathbf{E}_q^s(\mathbf{r}', \omega) + \mathbf{E}^b(\mathbf{r}', \omega) \right]d^3\mathbf{r}'. \quad (13)$$

According to Eq. (12), the scattered field $\boldsymbol{\Psi}^s$ of the system can be obtained as

$$\boldsymbol{\Psi}^s(\mathbf{r}, \omega) = \sum_{p=1}^{P} \boldsymbol{\Psi}_p^s(\mathbf{r}, \omega), \quad (14)$$

with $\boldsymbol{\Psi}_p^s(\mathbf{r}, \omega)$ being the scattered field of the pth MNR.

Comparing Eqs. (11) and (13), one can find that $\boldsymbol{\Psi}_p^s$ is just the field scattered by the pth MNR with $\sum_{q \neq p} \mathbf{E}_q^s(\mathbf{r}, \omega) + \mathbf{E}^b(\mathbf{r}, \omega)$ as the background field (i.e., incident field). This indicates that the coupling between MNRs in the system comes from the mutual excitation of the scattered field of each MNR, which is essentially the multiple-scattering picture used (but a demonstration of legitimacy not provided) in [51,52,59]. Then the scattered field of the pth MNR given by Eq. (13) can be expanded upon the basis of QNMs of the pth MNR via the QNM expansion formalism introduced in the previous section, so as to obtain the QNM coupling equations as stated below. Therefore, the Green's dyadic considerations here provide a rigorous demonstration of the multiple-scattering picture [Eqs. (12) and (13)], and provide a logic connection between this multiple-scattering picture and the QNM expansion formalism [Eqs. (6)–(8)].

According to Eq. (7), $\boldsymbol{\Psi}_p^s$ can be expanded upon M_p (depending on p) QNMs of the pth MNR,

$$\boldsymbol{\Psi}_p^s(\mathbf{r}, \omega) = \sum_{m=1}^{M_p} \alpha_{p,m}(\omega)\tilde{\boldsymbol{\Psi}}_{p,m}(\mathbf{r}), \tag{15}$$

where the mth-order QNM field $\tilde{\boldsymbol{\Psi}}_{p,m} = [\tilde{\mathbf{E}}_{p,m}, \tilde{\mathbf{H}}_{p,m}]$ of the pth MNR satisfies the source-free Maxwell's equations,

$$\nabla \times \tilde{\mathbf{E}}_{p,m} = i\tilde{\omega}_{p,m}\mu_0\tilde{\mathbf{H}}_{p,m},$$

$$\nabla \times \tilde{\mathbf{H}}_{p,m} = -i\tilde{\omega}_{p,m}\boldsymbol{\varepsilon}_p(\mathbf{r}, \tilde{\omega}_{p,m}) \cdot \tilde{\mathbf{E}}_{p,m}.$$

According to Eq. (8), the QNM expansion coefficient $\alpha_{p,m}$ in Eq. (15) can be expressed as

$$
\begin{aligned}
\alpha_{p,m}(\omega) &= \frac{-\omega \iiint_{V_p} \tilde{\mathbf{E}}_{p,m}(\mathbf{r}) \cdot \Delta\boldsymbol{\varepsilon}_p(\mathbf{r}, \omega) \cdot \left[\sum_{q \neq p} \mathbf{E}_q^s(\mathbf{r}, \omega) + \mathbf{E}^b(\mathbf{r}, \omega)\right] d^3\mathbf{r}}{(\omega - \tilde{\omega}_{p,m})F_{p,m}} \\
&= \frac{-\omega \iiint_{V_p} \tilde{\mathbf{E}}_{p,m}(\mathbf{r}) \cdot \Delta\boldsymbol{\varepsilon}_p(\mathbf{r}, \omega) \cdot \left[\sum_{q \neq p} \sum_{n=1}^{M_q} \alpha_{q,n}(\omega)\tilde{\mathbf{E}}_{q,n}(\mathbf{r}) + \mathbf{E}^b(\mathbf{r}, \omega)\right] d^3\mathbf{r}}{(\omega - \tilde{\omega}_{p,m})F_{p,m}}.
\end{aligned} \tag{16}
$$

In Eq. (16), V_p represents the region where the pth MNR is located, $F_{p,m}$ is the pseudoenergy of the mth-order QNM of the pth MNR expressed as

$$F_{p,m} = \iiint_{R^3} \left\{ \tilde{\mathbf{E}}_{p,m} \cdot \left. \frac{\partial[\omega\boldsymbol{\varepsilon}_p(\mathbf{r}, \omega)]}{\partial\omega} \right|_{\omega = \tilde{\omega}_{p,m}} \cdot \tilde{\mathbf{E}}_{p,m} - \mu_0\tilde{\mathbf{H}}_{p,m} \cdot \tilde{\mathbf{H}}_{p,m} \right\} d^3\mathbf{r},$$

and the second equality is obtained by using Eq. (15).

In the following, we will show that Eq. (16) can form a set of linear inhomogeneous equations with the QNM expansion coefficients $\alpha_{p,m}$ as the unknowns. For this purpose, the QNM-coupling coefficients are defined as

$$
\begin{aligned}
&\kappa_{(p,m),(q,n)}(\omega) \\
&= \frac{-\omega}{F_{p,m}} \iiint_{V_p} \tilde{\mathbf{E}}_{p,m}(\mathbf{r}) \cdot \Delta\boldsymbol{\varepsilon}_p(\mathbf{r}, \omega) \cdot \tilde{\mathbf{E}}_{q,n}(\mathbf{r}) d^3\mathbf{r}, \quad (17)
\end{aligned}
$$

with $p \neq q$. $\kappa_{(p,m),(q,n)}$ denotes the coefficient of the mth-order QNM of the pth MNR excited by the nth-order QNM of the qth MNR. It depends on the QNMs' field distribution and the relative position between the pth and the qth MNRs. The QNM excitation coefficients are defined as

$$\beta_{p,m}(\omega) = \frac{-\omega}{F_{p,m}} \iiint \tilde{\mathbf{E}}_{p,m}(\mathbf{r}) \cdot \Delta\boldsymbol{\varepsilon}_p(\mathbf{r}, \omega) \cdot \mathbf{E}^b(\mathbf{r}, \omega) d^3\mathbf{r}, \tag{18}$$

which denotes the coefficient of the mth-order QNM of the pth MNR excited by the background field $\mathbf{E}^b$. Then Eq. (16) can be rewritten as a set of linear inhomogeneous equations,

$$
\begin{aligned}
\beta_{p,m} &= (\omega - \tilde{\omega}_{p,m})\alpha_{p,m} - \sum_{q \neq p} \sum_{n=1}^{M_q} \kappa_{(p,m),(q,n)}\alpha_{q,n} \\
&= \omega\alpha_{p,m} - \sum_{q=1}^{P} \sum_{n=1}^{M_q} [\delta_{p,q}\delta_{m,n}\tilde{\omega}_{p,m} \\
&\quad + (1 - \delta_{p,q})\kappa_{(p,m),(q,n)}]\alpha_{q,n}, \tag{19}
\end{aligned}
$$

with $\delta_{p,q}$ and $\delta_{m,n}$ being the Kronecker delta. Equation (19) can be rewritten in a concise matrix form,

$$[\omega\mathbf{I} - \mathbf{K}(\omega)]\mathbf{a} = \mathbf{b}(\omega), \tag{20}$$

where $\mathbf{a}$ is a column vector with the QNM expansion coefficient $\alpha_{p,m}$ as its (p, m)th element, the frequency-dependent $\mathbf{b}(\omega)$ is a column vector with the QNM excitation coefficient $\beta_{p,m}$ as its (p, m)th element, and $\mathbf{I}$ represents the identity matrix. The element of the (p, m)th row and (q, n)th column of the frequency-dependent matrix $\mathbf{K}(\omega)$ is the QNM coupling coefficient $\kappa_{(p,m),(q,n)}$ for $p \neq q$, and is $\delta_{m,n}\tilde{\omega}_{p,m}$ for $p = q$.

Solving Eq. (20), one then obtains the QNM expansion coefficients $\alpha_{p,m}$, for which the $\mathbf{b}$ and $\mathbf{K}$ are determined by the given background field and the already solved QNMs of each MNR in the coupled-MNR system. By substituting the solved $\alpha_{p,m}$ into Eqs. (15) and (14), the scattered field of the system can be obtained. Therefore, a repetition of solving Maxwell's equations is not needed when changing the number and the relative position of MNRs, or the frequency and spatial distribution of the background field. In particular, if the scattered field of a single MNR can be expanded upon a small set of QNMs (which is often true such as for the example in Fig. 1), the computational amount of calculating the scattered field of the coupled-MNRs system can be greatly reduced by using the present QNM coupling theory. Besides, the present theory shows that the coupling between MNRs originates from the mutual excitation of the scattered fields of each MNR [see Eq. (13)], which provides an intuitive physical picture for understanding the coupling mechanism of coupled MNRs.

B. QNM coupling theory for the QNMs of the coupled-MNR system

Here we use $\tilde{\boldsymbol{\Psi}}_r(\mathbf{r})$ to denote the field of the rth-order QNM of the coupled-MNR system. By substituting Eq. (15) into Eq. (14), $\tilde{\boldsymbol{\Psi}}_r(\mathbf{r})$ can then be expanded in terms of QNMs of each MNR as

$$\tilde{\boldsymbol{\Psi}}_r(\mathbf{r}) = \tilde{\boldsymbol{\Phi}}^T(\mathbf{r})\mathbf{a}, \tag{21}$$

where $\tilde{\boldsymbol{\Phi}}^T(\mathbf{r})$ is a row vector with $\tilde{\boldsymbol{\Psi}}_{p,m}(\mathbf{r})$ as its (p, m)th element. For QNMs of the coupled MNRs, the background field $\boldsymbol{\Psi}^b$ is absent. Thus, the QNM coupling Eq. (20) becomes

a set of linear homogeneous equations after removing the inhomogeneous term $\mathbf{b}$,

$$[\omega \mathbf{I} - \mathbf{K}(\omega)]\mathbf{a} = \mathbf{0}. \tag{22}$$

The coefficient $\mathbf{a}$ is thus given by the nontrivial solution of Eq. (22), whose existence requires

$$\det[\omega \mathbf{I} - \mathbf{K}(\omega)] = 0. \tag{23}$$

Equation (23) forms a transcendental equation of frequency ω, and its solution is just the complex eigenfrequency $\omega = \tilde{\omega}_r$ of the coupled-MNR system. Substituting the solved eigenfrequency $\omega = \tilde{\omega}_r$ into Eq. (22), one can then solve its nontrivial solution $\mathbf{a} = \tilde{\mathbf{a}}_r$, whose (p, m)th element is denoted by $\tilde{\alpha}_{r,(p,m)}$.

For a system with an arbitrary frequency dispersion of permittivity, the transcendental Eq. (23) can be solved with an iterative algorithm, such as the linear interpolation iterative algorithm [75]. For an isotropic medium with a dispersive permittivity described by an N-pole Drude-Lorentz model [55], Eq. (22) can reduce to a polynomial matrix eigenvalue problem. For example, for a system in which the background medium is nondispersive and the relative permittivity of MNRs is described by the Drude model, the permittivity change of the pth MNR is

$$\Delta \varepsilon_p(\omega) = \varepsilon_0 \big[\varepsilon_{\infty,p} - \omega_{0,p}^2 / (\omega^2 + i\omega\gamma_p) \big] - \varepsilon_b,$$

where $\varepsilon_{\infty,p}$, $\omega_{0,p}$, and γ_p are real-valued parameters of the Drude model, and ε_b is the background permittivity. With the $\Delta \varepsilon_p(\omega)$ inserted into, Eq. (22) can then be rewritten as a polynomial eigenvalue problem,

$$\omega^2 \mathbf{C}_2 \mathbf{a} + \omega \mathbf{C}_1 \mathbf{a} + \mathbf{C}_0 \mathbf{a} = \mathbf{0}. \tag{24}$$

For $p \neq q$, the element of the (p, m)th row and (q, n)th column of matrices $\mathbf{C}_2$, $\mathbf{C}_1$, and $\mathbf{C}_0$ are $(\varepsilon_0 \varepsilon_{\infty,p} - \varepsilon_b)M_{(p,m),(q,n)}$, $i\gamma_p(\varepsilon_0 \varepsilon_{\infty,p} - \varepsilon_b)M_{(p,m),(q,n)}$, and $-\varepsilon_0\omega_{0,p}^2 M_{(p,m),(q,n)}$, respectively, with $M_{(p,m),(q,n)} = \iiint_{V_p} \tilde{\mathbf{E}}_{p,m}(\mathbf{r}) \cdot \tilde{\mathbf{E}}_{q,n}(\mathbf{r}) d^3\mathbf{r}$. For $p = q$, the element of the (p, m)th row and (p, n)th column of $\mathbf{C}_2$, $\mathbf{C}_1$, and $\mathbf{C}_0$ are $\delta_{m,n}F_{p,m}$, $\delta_{m,n}(i\gamma_p - \tilde{\omega}_{p,m})F_{p,m}$ and $-i\delta_{m,n}\tilde{\omega}_{p,m}\gamma_p F_{p,m}$, respectively. Details on the derivation of Eq. (24) are provided in Supplemental Material Sec. S2 [69]. Effective algorithms have been developed for solving the polynomial eigenvalue problems such as Eq. (24) [87,88].

Compared with the classical temporal coupled-mode theory (TCMT) [28,38], our theory has a more general applicability and can provide a rigorous electromagnetic foundation for the former, as explained below. Removing the inhomogeneous term $\beta_{p,m}$ and performing a Fourier transformation of both sides of Eq. (19), one then obtains the time-domain QNM coupling equations,

$$i\frac{dA_{p,m}(t)}{dt} = \tilde{\omega}_{p,m} A_{p,m}(t)$$
$$+ \sum_{q \neq p} \sum_{n=1}^{M_q} \mathcal{F}_\omega\{\kappa_{(p,m),(q,n)}(\omega)\}_t * A_{q,n}(t), \tag{25}$$

with

$$A_{p,m}(t) = \mathcal{F}_\omega\{\alpha_{p,m}(\omega)\}_t = \frac{1}{2\pi} \int_{-\infty}^{\infty} \alpha_{p,m}(\omega) \exp(-i\omega t) d\omega$$

being the QNM expansion coefficient in the time domain, and $*$ denoting the convolution. For the case of weak coupling, the frequency shift $\Delta\omega = \omega - \tilde{\omega}_{p,m}$ (with $\omega = \tilde{\omega}_r$) caused by coupling will be small according to our theory. Then the coupling coefficient $\kappa_{(p,m),(q,n)}(\omega) \approx \kappa_{(p,m),(q,n)}(\tilde{\omega}_{p,m})$ becomes approximately independent of the frequency ω since $\kappa_{(p,m)(q,n)}(\omega)$ is a slowly varying function of ω. Then there is

$$\mathcal{F}_\omega\{\kappa_{(p,m),(q,n)}(\omega)\}_t * A_{q,n}(t) \approx \kappa_{(p,m),(q,n)}(\tilde{\omega}_{p,m})A_{q,n}(t),$$

so that Eq. (25) reduces to the form of the classical TCMT. In addition, for the classical TCMT, the electric field in the integral expression of the coupling coefficient $\kappa_{(p,m),(q,n)}$ takes a complex conjugate [e.g., Eq. (2.16) in Ref. [28]], which is rigorously correct only for lossless media. Thus, the classical TCMT is a special form of the present QNM coupling theory for the case of weak coupling and low loss. In some cases, TCMT is used as a phenomenological theory in which the parameters (such as the coupling coefficient $\kappa_{(p,m)(q,n)}$) are obtained by fitting full-wave numerical results or experimental data [27,40,41].

In Supplemental Material Sec. S3 [69], we provide a logic crosscheck on the validity of our theory from the viewpoint of the original Maxwell's equations instead of the Green's-function formalism, which will show more details concerning the relation between our theory and other theories as provided in Supplemental Material Sec. S4 [69].

C. QNM coupling theory for the QNM expansion of the scattered field of the coupled-MNR system

In this section, we will take the QNMs of the coupled MNRs [given by Eqs. (21)–(23)] as a set of basis to expand the scattered field of the coupled MNRs and build up the QNM expansion theory of the coupled MNRs, so as to explicitly show the impact of QNMs of the coupled-MNR system on the resonance properties of the scattered field of the system.

For this purpose, the solution of the inhomogeneous Eq. (20) can be expressed as

$$\mathbf{a} = \frac{[\omega \mathbf{I} - \mathbf{K}(\omega)]^* \mathbf{b}(\omega)}{\det[\omega \mathbf{I} - \mathbf{K}(\omega)]}. \tag{26}$$

Equation (26) can be obtained with the only assumption that the matrix $\omega \mathbf{I} - \mathbf{K}(\omega)$ is invertible, which yields $[\omega \mathbf{I} - \mathbf{K}(\omega)]^{-1} = \frac{[\omega \mathbf{I} - \mathbf{K}(\omega)]^*}{\det[\omega \mathbf{I} - \mathbf{K}(\omega)]}$ with $[\omega \mathbf{I} - \mathbf{K}(\omega)]^*$ denoting the adjugate matrix of $\omega \mathbf{I} - \mathbf{K}(\omega)$. Equation (26) shows that at the complex eigenfrequency $\omega = \tilde{\omega}_r$ of the QNMs of the coupled MNRs, the coefficient $\mathbf{a}$ of QNMs of each MNR tends to infinity in view of $\det[\tilde{\omega}_r \mathbf{I} - \mathbf{K}(\tilde{\omega}_r)] = 0$, so that the scattered field $\mathbf{\Psi}^s(\mathbf{r}, \omega)$ given by Eqs. (14) and (15) also tends to infinity. Therefore, $\tilde{\omega}_r$ is the complex frequency pole of $\mathbf{\Psi}^s(\mathbf{r}, \omega)$. Then by assuming that $\mathbf{\Psi}^s(\mathbf{r}, \omega)$ is a meromorphic function of ω, and is bounded as $\omega \to \infty$ in the complex plane [the latter can be replaced by some more sophisticated and exactly specified condition such as that before Eq. (7.20) in [89]], and that $\tilde{\omega}_r$ is the first-order pole of $\mathbf{\Psi}^s(\mathbf{r}, \omega)$ (i.e., there is no QNM degenerate exceptional point [26,27]) with $\tilde{\omega}_r \neq 0$, $\mathbf{\Psi}^s(\mathbf{r}, \omega)$ can be expressed via the complex pole expansion

theorem (Mittag-Leffler theorem) [48,89,90] as

$$\Psi^s(\mathbf{r}, \omega) = \Psi^s(\mathbf{r}, 0) + \sum_r \frac{\omega/\tilde{\omega}_r}{\omega - \tilde{\omega}_r} \mathbf{P}_r(\mathbf{r}). \qquad (27)$$

In Eq. (27), $\mathbf{P}_r(\mathbf{r}) = \lim_{\omega \to \tilde{\omega}_r}[(\omega - \tilde{\omega}_r)\Psi^s(\mathbf{r}, \omega)]$ is the residue of $\Psi^s(\mathbf{r}, \omega)$ at the complex pole $\tilde{\omega}_r$, and there is $\Psi^s(\mathbf{r}, 0) = 0$ in view that the scattering effect of the finite-size resonator will vanish for an infinitely large wavelength. To calculate $\mathbf{P}_r(\mathbf{r})$, one can substitute Eq. (26) into Eqs. (15) and (14) so as to express $\Psi^s(\mathbf{r}, \omega)$ in a matrix form,

$$\Psi^s(\mathbf{r}, \omega) = \frac{\tilde{\boldsymbol{\Phi}}^T(\mathbf{r})[\omega\mathbf{I} - \mathbf{K}(\omega)]^*\mathbf{b}(\omega)}{\det[\omega\mathbf{I} - \mathbf{K}(\omega)]}, \qquad (28)$$

where $\tilde{\boldsymbol{\Phi}}^T(\mathbf{r})$ is the row vector defined in Eq. (21), with its (p, m)th element being the mth-order QNM field $\tilde{\Psi}_{p,m}(\mathbf{r})$ of the pth MNR. With Eq. (28), one can obtain

$$\mathbf{P}_r(\mathbf{r}) = \frac{\tilde{\boldsymbol{\Phi}}^T(\mathbf{r})[\tilde{\omega}_r\mathbf{I} - \mathbf{K}(\tilde{\omega}_r)]^*\mathbf{b}(\tilde{\omega}_r)}{\frac{\partial \det[\omega\mathbf{I} - \mathbf{K}(\omega)]}{\partial \omega}\big|_{\omega=\tilde{\omega}_r}}. \qquad (29)$$

Assuming that there is no QNM degeneracy, that is, there is only one linearly independent solution for Eq. (22) at $\omega = \tilde{\omega}_r$, one can obtain $\text{rank}[\tilde{\omega}_r\mathbf{I} - \mathbf{K}(\tilde{\omega}_r)] = \dim[\tilde{\omega}_r\mathbf{I} - \mathbf{K}(\tilde{\omega}_r)] - 1$, and $\text{rank}\{[\omega_r\mathbf{I} - \mathbf{K}(\omega_r)]^*\} = 1$, where dim means the dimension of a square matrix. Then $[\omega_r\mathbf{I} - \mathbf{K}(\omega_r)]^*$ can be expressed as [91]

$$[\tilde{\omega}_r\mathbf{I} - \mathbf{K}(\tilde{\omega}_r)]^* = \tilde{\mathbf{a}}_r\mathbf{c}_r^T, \qquad (30)$$

where $\mathbf{c}_r^T$ is a row vector and $\tilde{\mathbf{a}}_r$ is the nontrivial solution of Eq. (22). By successively substituting Eqs. (30) and (21) of $\tilde{\Psi}_r(\mathbf{r}) = \tilde{\boldsymbol{\Phi}}^T(\mathbf{r})\mathbf{a}_r$ into Eq. (29), and then substituting Eq. (29) into Eq. (27), the scattered field $\Psi^s(\mathbf{r}, \omega)$ can be finally expanded in terms of QNM field $\tilde{\Psi}_r(\mathbf{r})$ of the coupled MNRs,

$$\Psi^s(\mathbf{r}, \omega) = \sum_r \frac{\omega/\tilde{\omega}_r}{\omega - \tilde{\omega}_r} \frac{\mathbf{c}_r^T\mathbf{b}(\tilde{\omega}_r)}{\frac{\partial \det[\omega\mathbf{I} - \mathbf{K}(\omega)]}{\partial \omega}\big|_{\omega=\tilde{\omega}_r}} \tilde{\Psi}_r(\mathbf{r}). \qquad (31)$$

Equation (31) explicitly provides the influence of QNMs of the coupled-MNR system on the resonance properties of the scattered field of the system, and the influence of the QNM excitation coefficient $\beta_{p,m}$, coupling coefficient $\kappa_{(p,m),(q,n)}$ and complex eigenfrequency $\tilde{\omega}_{p,m}$ of MNRs [see the definitions of $\mathbf{b}$ and $\mathbf{K}$ in Eq. (20)] on the expansion coefficient upon the basis of the QNMs of the coupled MNRs.

Another way to derive the source-excitation coefficient of the hybridized QNM of the coupled-MNR system is to use the QNM expansion formalism in Sec. II with the hybridized QNMs as the expansion basis, as has been done in [52]. But this requires a normalization condition of the hybridized QNM which has not been formally confirmed [52]. This normalization condition is not required by Eq. (31), which is due to the use of the Mittag-Leffler theorem to derive the expansion of the source-excited field upon the basis of the hybridized QNMs. But note that for Eq. (31), this normalization condition [or equivalently, the calculation of the pseudoenergy $F_{p,m}$ defined in Eq. (16)] is still required for the QNMs of each single MNR.

D. Special case of two coupled QNMs

To gain a more concrete understanding of the above theory, in this section we will discuss a special case that frequently appears in practice, where the coupled-MNR system consists of two MNRs (referred to as an MNR dimer in this paper) [15,17,80,92], and only one QNM is considered for each MNR. It is assumed that the QNM considered for each MNR has a high Q factor and plays a decisive role in the resonance characteristics of the MNR, such as the numerical examples to be shown later in Figs. 4(a)–4(c). For this case, the scattered field can be expressed according to Eqs. (14) and (15) as

$$\Psi^s(\mathbf{r}, \omega) = \alpha_{1,1}(\omega)\tilde{\Psi}_{1,1}(\mathbf{r}) + \alpha_{2,1}(\omega)\tilde{\Psi}_{2,1}(\mathbf{r}), \qquad (32)$$

where the QNM expansion coefficients $\alpha_{1,1}$ and $\alpha_{1,2}$ satisfy the linear inhomogeneous Eq. (19),

$$(\omega - \tilde{\omega}_{1,1})\alpha_{1,1}(\omega) - \kappa_{(1,1),(2,1)}(\omega)\alpha_{2,1}(\omega) = \beta_{1,1}(\omega),$$
$$(\omega - \tilde{\omega}_{2,1})\alpha_{2,1}(\omega) - \kappa_{(2,1),(1,1)}(\omega)\alpha_{1,1}(\omega) = \beta_{2,1}(\omega). \quad (33)$$

Solving Eq. (33), one can obtain $\mathbf{a} = [\alpha_{1,1}, \alpha_{2,1}]^T$. Substituting it into Eq. (32), one then obtains the scattered field of the system.

To solve the QNMs of the MNR dimer, one can remove the inhomogeneous terms on the right-hand side of Eq. (33) and then obtain a set of linear homogeneous equations [i.e., Eq. (22)],

$$(\omega - \tilde{\omega}_{1,1})\alpha_{1,1}(\omega) - \kappa_{(1,1),(2,1)}(\omega)\alpha_{2,1}(\omega) = 0,$$
$$-\kappa_{(2,1),(1,1)}(\omega)\alpha_{1,1}(\omega) + (\omega - \tilde{\omega}_{2,1})\alpha_{2,1}(\omega) = 0. \quad (34)$$

The QNMs of the system are the nontrivial solutions of Eq. (34), whose existence requires the determinant of the coefficient matrix to be zero [i.e., Eq. (23)]. This yields that the complex resonance frequencies $\omega = \tilde{\omega}_r$ of the system satisfies

$$\omega = \frac{\tilde{\omega}_{1,1} + \tilde{\omega}_{2,1}}{2}$$
$$\pm \sqrt{\left(\frac{\tilde{\omega}_{1,1} - \tilde{\omega}_{2,1}}{2}\right)^2 + \kappa_{(1,1),(2,1)}(\omega)\kappa_{(2,1),(1,1)}(\omega)}. \quad (35)$$

Note that Eq. (35) is similar to the conclusion of TCMT [see Eq. (2.8) in [28]]. But differently, the coupling coefficient $\kappa_{(p,m),(q,n)} = \kappa$ in [28] is independent of frequency ω, which is only applicable to the case of weak coupling as explained earlier following Eq. (25). Besides, the electric field in the integral expression of κ takes a complex conjugate [see Eq. (2.16) in [28]], which is rigorously correct only for a lossless medium.

Equation (35) forms a transcendental equation of ω since its right-hand side is a function of ω [denoted by $f(\omega)$]. However, since $f(\omega)$ is a slowly varying function of ω in view that $\kappa_{(p,m),(q,n)}$ is a slowly varying function of ω [see Eq. (17)], the transcendental Eq. (35) can be solved by using the contractive mapping method [93] with the iterative formula $\omega_{N+1} = f(\omega_N)$, $N = 0, 1, \dots$. This method converges fast and is insensitive to the initial value ω_0 of the iteration. In fact, if $f(\omega)$ is approximately treated as a constant, Eq. (35) simply provides an analytical expression for the complex resonant frequency $\tilde{\omega}_r$. Therefore, commonly there are only two
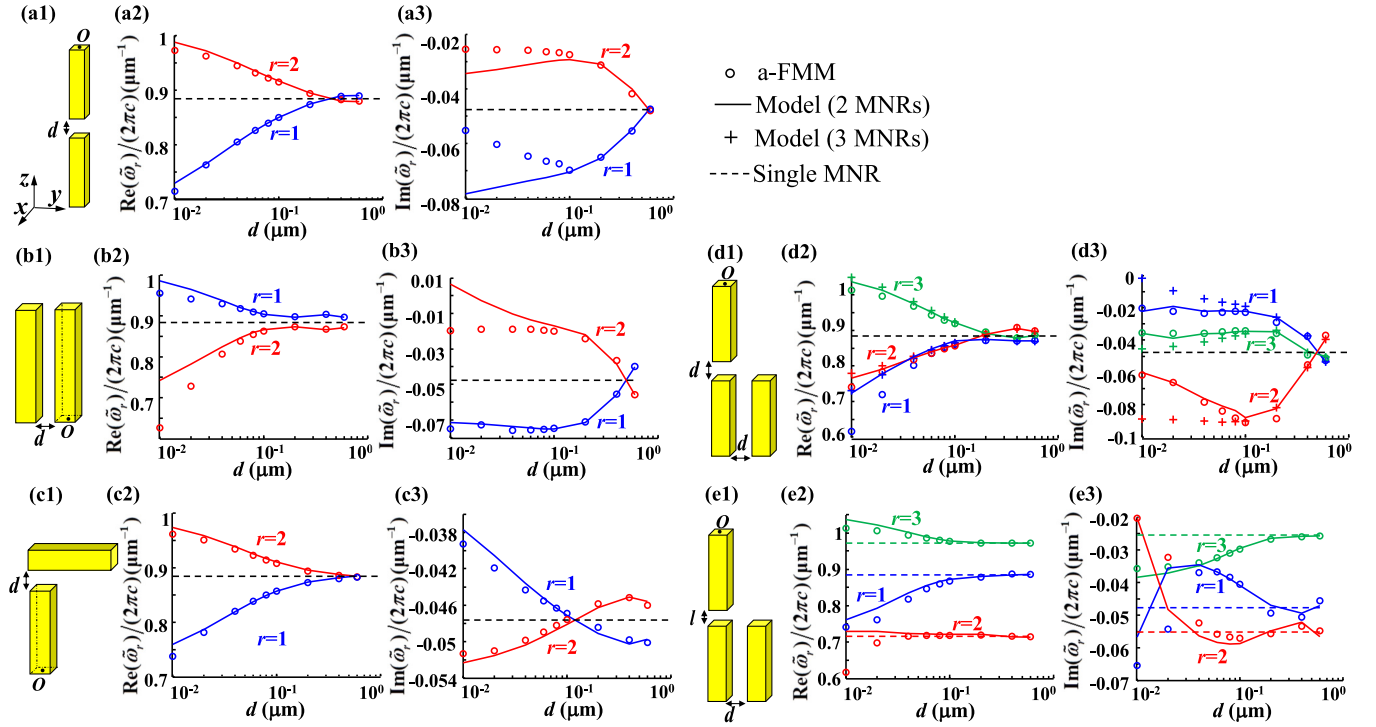
FIG. 4. (a1)–(e1) Schematic diagrams of the five coupled MNRs. The structural parameters of the single-wire MNR are the same as those in Fig. 1(a). There is $l = 10$ nm in (e1). The origin O of the coordinate system is located at the center of the antenna end face. (a2)–(e2) and (a3)–(e3) respectively show the real and imaginary parts of the complex resonance frequency $\tilde{\omega}_r$ of the coupled MNRs as functions of the distance d. Curves with different colors show the results of different orders of QNMs [i.e., with different r, $r = 1$ and 2 in (a)–(c) corresponding to symmetric and antisymmetric QNMs, respectively], which are obtained with full-wave a-FMM numerical calculation (circles) and the QNM coupling theory (solid curves and pluses). In (d) and (e), the solid curves and pluses show the predictions of the model in which the trimer antenna is treated to be composed of a dipole antenna [as sketched in (a1)] and a single-wire antenna (i.e., 2 MNRs), or of three single-wire antennas (i.e., 3 MNRs), respectively.

solutions of QNMs with high Q factors given by Eq. (35) that are close to $\tilde{\omega}_{1,1}$ and $\tilde{\omega}_{2,1}$, which correspond to the $+$ and $-$ on the right side of Eq. (35). The other solutions are far away from $\tilde{\omega}_{1,1}$ and $\tilde{\omega}_{2,1}$, which either have low Q factors or locate out of the spectral range of interest, and thus can be neglected in the QNM expansion of the scattered field $\mathbf{\Psi}^s$ [see Eq. (40) below]. This will be verified by the numerical examples in Figs. 4(a)–4(c) and Figs. 6(a)–6(c), to be shown later.

According to Eq. (31), the scattered field of the MNR dimer can be expanded upon the basis of QNMs of the MNR dimer given by Eq. (34). The denominator in the right side of Eq. (31) can be simplified to be

$$\frac{\partial \det[\omega \mathbf{I} - \mathbf{K}(\omega)]}{\partial \omega}\bigg|_{\omega = \tilde{\omega}_r}$$
$$= 2\tilde{\omega}_r - \tilde{\omega}_{1,1} - \tilde{\omega}_{2,1} - \frac{\partial(\kappa_{(1,1)(2,1)}\kappa_{(2,1)(1,1)})}{\partial \omega}\bigg|_{\omega = \tilde{\omega}_r}. \quad (36)$$

The analytical expression of $\mathbf{c}_r$ in Eq. (31) will be derived below. For the QNM with a complex resonance frequency $\omega = \tilde{\omega}_r$, the nontrivial solution of Eq. (34) is

$$\mathbf{a} = \tilde{\mathbf{a}}_r = [\tilde{\alpha}_{r,(1,1)}, \tilde{\alpha}_{r,(2,1)}]^T$$
$$= [1, \kappa_{(2,1),(1,1)}(\tilde{\omega}_r)/(\tilde{\omega}_r - \tilde{\omega}_{2,1})]^T$$
$$= [1, (\tilde{\omega}_r - \tilde{\omega}_{1,1})/\kappa_{(1,1),(2,1)}(\tilde{\omega}_r)]^T, \quad (37)$$

where the normalization $\tilde{\alpha}_{r,(1,1)} = 1$ is specified for the QNM. For the MNR dimer, one can obtain

$$[\tilde{\omega}_r \mathbf{I} - \mathbf{K}(\tilde{\omega}_r)]^* = \begin{bmatrix} \tilde{\omega}_r - \tilde{\omega}_{2,1} & \kappa_{(1,1),(2,1)}(\tilde{\omega}_r) \\ \kappa_{(2,1),(1,1)}(\tilde{\omega}_r) & \tilde{\omega}_r - \tilde{\omega}_{1,1} \end{bmatrix}. \quad (38)$$

Substituting Eqs. (37) and (38) into Eq. (30), we then obtain the expression of $\mathbf{c}_r$ as

$$\mathbf{c}_r = [\tilde{\omega}_r - \tilde{\omega}_{2,1}, \kappa_{(1,1)(2,1)}(\tilde{\omega}_r)]^T. \quad (39)$$

By substituting Eqs. (36) and (39) into Eq. (31), the scattered field of the MNR dimer can be finally expressed as

$$\mathbf{\Psi}^s(\mathbf{r}, \omega) = \sum_r \tilde{\mathbf{\Psi}}_r(\mathbf{r}) \frac{\omega/\tilde{\omega}_r}{\omega - \tilde{\omega}_r}$$
$$\times \left[\frac{(\omega - \tilde{\omega}_{2,1})\beta_{1,1} + \kappa_{(1,1)(2,1)}\beta_{2,1}}{2\omega - \tilde{\omega}_{1,1} - \tilde{\omega}_{2,1} - \frac{\partial(\kappa_{(1,1)(2,1)}\kappa_{(2,1)(1,1)})}{\partial \omega}} \right]_{\omega = \tilde{\omega}_r}, \quad (40)$$

where the QNM field $\tilde{\mathbf{\Psi}}_r(\mathbf{r})$ of the MNR dimer can be obtained by substituting Eq. (37) into Eq. (32) as [i.e., Eq. (21)]

$$\tilde{\mathbf{\Psi}}_r(\mathbf{r}) = \tilde{\mathbf{\Psi}}_{1,1}(\mathbf{r}) + \kappa_{(2,1),(1,1)}(\tilde{\omega}_r)\tilde{\mathbf{\Psi}}_{2,1}(\mathbf{r})/(\tilde{\omega}_r - \tilde{\omega}_{2,1}). \quad (41)$$

Next we will consider a more special case that two identical MNRs are placed parallel to each other with a mirror symmetry, such as the structures shown in Figs. 4(a1)–4(b1).

Here it is assumed that the background medium and the medium of the MNRs are isotropic and uniform. For the pth MNR ($p = 1, 2$), its permittivity distribution can be denoted by $\boldsymbol{\varepsilon}_p(\mathbf{r}, \omega) = \varepsilon^0(\mathbf{r} - \mathbf{r}_p, \omega)$ (with $\mathbf{r}_p$ being the central position vector of the pth MNR), and the complex resonance frequency and field of its supported QNM are denoted by $\tilde{\omega}_{p,1} = \tilde{\omega}_0$ and $\tilde{\boldsymbol{\Psi}}_{p,1}(\mathbf{r}) = \tilde{\boldsymbol{\Psi}}_0(\mathbf{r} - \mathbf{r}_p)$, respectively. Then according to Eq. (17), the QNM coupling coefficient can be expressed as

$$\kappa_{(1,1),(2,1)}(\omega) = \kappa_{(2,1),(1,1)}(\omega) = \kappa(\omega) = \frac{-\omega \Delta \varepsilon^0(\omega) M_0}{F_0},$$

$$\tag{42}$$

with

$$F_0 = F_{1,1} = F_{2,1} = \iiint_{R^3}$$
$$\times \left\{ \tilde{\mathbf{E}}_0(\mathbf{r}) \cdot \frac{\partial(\omega \boldsymbol{\varepsilon}^0)}{\partial \omega} \bigg|_{\omega = \tilde{\omega}_0} \cdot \tilde{\mathbf{E}}_0(\mathbf{r}) - \mu_0 \tilde{\mathbf{H}}_0(\mathbf{r}) \cdot \tilde{\mathbf{H}}_0(\mathbf{r}) \right\} d^3\mathbf{r},$$

and

$$M_0 = M_{(1,1),(2,1)} = M_{(2,1),(1,1)}$$
$$= \iiint_{V_p} \tilde{\mathbf{E}}_0(\mathbf{r} - \mathbf{r}_2) \cdot \tilde{\mathbf{E}}_0(\mathbf{r} - \mathbf{r}_1) d^3\mathbf{r}.$$

Here the $F_{p,m}$ and $M_{(p,m),(q,n)}$ are first defined following Eqs. (16) and (24), respectively. The V_p denotes the region of the pth MNR, and $\Delta \varepsilon^0 = \varepsilon^0 - \varepsilon_b$ is the permittivity change of the pth MNR with ε_b being the permittivity of the background medium. Thus, Eq. (35) can reduce to

$$\omega = \tilde{\omega}_0 \pm \kappa(\omega). \tag{43}$$

Substituting Eq. (43) into Eq. (34), one can obtain the expansion coefficients $\mathbf{a} = \tilde{\mathbf{a}}_+$ and $\mathbf{a} = \tilde{\mathbf{a}}_-$ [corresponding to $\tilde{\omega}_0 + \kappa(\omega)$ and $\tilde{\omega}_0 - \kappa(\omega)$ in Eq. (43), respectively] of the two types of QNMs supported by the system as

$$\alpha_{1,1} = \alpha_{2,1}, \quad \alpha_{1,1} = -\alpha_{2,1}, \tag{44}$$

which are termed symmetric and antisymmetric QNMs, respectively. With the normalization of $\alpha_{1,1} = 1$, the fields of QNMs corresponding to Eq. (44) can be expressed as

$$\tilde{\boldsymbol{\Psi}}_+(\mathbf{r}) = \tilde{\boldsymbol{\Psi}}_0(\mathbf{r} - \mathbf{r}_1) + \tilde{\boldsymbol{\Psi}}_0(\mathbf{r} - \mathbf{r}_2),$$
$$\tilde{\boldsymbol{\Psi}}_-(\mathbf{r}) = \tilde{\boldsymbol{\Psi}}_0(\mathbf{r} - \mathbf{r}_1) - \tilde{\boldsymbol{\Psi}}_0(\mathbf{r} - \mathbf{r}_2). \tag{45}$$

With Eq. (39), one can obtain

$$\mathbf{c}_r = \mathbf{c}_+ = \kappa(\tilde{\omega}_+)[1, 1]^T = \kappa(\tilde{\omega}_+)\tilde{\mathbf{a}}_+,$$

and

$$\mathbf{c}_r = \mathbf{c}_- = -\kappa(\tilde{\omega}_-)[1, -1]^T = -\kappa(\tilde{\omega}_-)\tilde{\mathbf{a}}_-,$$

for symmetric and antisymmetric modes, respectively. Thus, Eq. (40) can be simplified to be

$$\boldsymbol{\Psi}^s(\mathbf{r}, \omega) = \frac{\omega/\tilde{\omega}_+}{\omega - \tilde{\omega}_+} \left[\frac{\beta_{1,1} + \beta_{2,1}}{2(1 - \partial\kappa/\partial\omega)} \right]_{\omega = \tilde{\omega}_+} \tilde{\boldsymbol{\Psi}}_+(\mathbf{r})$$
$$+ \frac{\omega/\tilde{\omega}_-}{\omega - \tilde{\omega}_-} \left[\frac{\beta_{1,1} - \beta_{2,1}}{2(1 + \partial\kappa/\partial\omega)} \right]_{\omega = \tilde{\omega}_-} \tilde{\boldsymbol{\Psi}}_-(\mathbf{r}). \tag{46}$$

The QNM expansion of Eq. (46) contains only two terms due to the reason mentioned after Eq. (35). Equation (46) indicates that if the spatial distribution of the excitation source satisfies a certain symmetry, then only symmetric or antisymmetric mode can be excited. For example, for $\beta_{1,1} = \beta_{2,1}$, the excitation coefficient of $\tilde{\boldsymbol{\Psi}}_-(\mathbf{r})$ becomes zero so that the antisymmetric mode behaves as a dark mode [24].

Equation (43) indicates that the shifts of eigenfrequencies $\tilde{\omega}_\pm$ of the MNR dimer relative to the eigenfrequency $\tilde{\omega}_0$ of a single MNR are $\Delta\tilde{\omega}_\pm = \tilde{\omega}_\pm - \tilde{\omega}_0 = \pm\kappa(\tilde{\omega}_\pm)$. Since κ is a function of ω, there is $\Delta\tilde{\omega}_+ \neq \Delta\tilde{\omega}_-$ in general. In the following, we will provide an analysis of the dependence of κ on ω with the model, so as to compare the values of $|\text{Re}(\Delta\tilde{\omega}_+)| = |\text{Re}[\kappa(\tilde{\omega}_+)]|$ and $|\text{Re}[\Delta(\tilde{\omega}_-)]| = |\text{Re}[\kappa(\tilde{\omega}_-)]|$. We only discuss the QNMs with high Q factors and assume $\tilde{\omega}_\pm \approx \text{Re}(\tilde{\omega}_\pm)$, i.e., ω is approximately of real value in the following discussion. The frequency dispersion of permittivity is assumed to be described by the Drude model. Then the scatterer permittivity change is

$$\Delta\varepsilon^0(\omega) = \varepsilon_0\varepsilon_\infty - \frac{\varepsilon_0\omega_0^2}{\omega^2 + i\omega\gamma} - \varepsilon_b.$$

Substituting it into Eq. (42), one can obtain

$$|\text{Re}[\kappa(\omega)]| = \left| \text{Re}\left[\frac{M_0}{F_0}(\varepsilon_0\varepsilon_\infty - \varepsilon_b)\omega \right] - \text{Re}\left[\frac{M_0}{F_0} \frac{\varepsilon_0\omega_0^2}{\omega + i\gamma} \right] \right|.$$

There is $(\varepsilon_0\varepsilon_\infty - \varepsilon_b)\omega < \varepsilon_0\omega_0^2/\omega$ in view of $\omega > 0$ and $\text{Re}[\Delta\varepsilon^0(\omega)] = (\varepsilon_0\varepsilon_\infty - \varepsilon_b - \frac{\varepsilon_0\omega_0^2}{\omega^2 + \gamma^2}) < 0$. Then neglecting γ in view of $\gamma \ll \omega$, one can obtain

$$|\text{Re}[\kappa(\omega)]| \approx \left| \text{Re}\left(\frac{M_0}{F_0} \right) \right| \left[\frac{\varepsilon_0\omega_0^2}{\omega} - (\varepsilon_0\varepsilon_\infty - \varepsilon_b)\omega \right]. \tag{47}$$

Equation (47) shows that the frequency dispersion of the permittivity causes $|\text{Re}[\kappa(\omega)]|$ to be a monotonically decreasing function of ω. Therefore, for $\text{Re}(\tilde{\omega}_+) > \text{Re}(\tilde{\omega}_-)$, there is $|\text{Re}[\kappa(\tilde{\omega}_+)]| < |\text{Re}[\kappa(\tilde{\omega}_-)]|$, which further yields $|\text{Re}(\Delta\tilde{\omega}_+)| < |\text{Re}(\Delta\tilde{\omega}_-)|$. This conclusion is consistent with the numerical results to be shown in Figs. 4(a) and 4(b) and with the reported experimental phenomena [94] and theoretical prediction [29].

IV. NUMERICAL TEST OF THE QNM COUPLING THEORY

To test the validity of the QNM coupling theory developed in this paper, it is used to predict the electromagnetic response of several representative coupled-MNR systems in comparison to the full-wave a-FMM numerical results. Five coupled-MNR systems are selected as numerical examples, including MNR dimers [80,95] and MNR trimers [11,96,97]. In the following Secs. IV A and IV B, the validity of the theory proposed in Sec. III B in predicting the complex resonance frequencies and the fields of the QNMs of the coupled-MNR system will be tested, respectively. In Sec. IV C, the validity of the theories proposed in Secs. III A and III C in predicting the source-excited scattered field of the coupled-MNR system will be tested. This test also shows the use of the theory for the design of antenna devices to achieve a broadband enhancement of the spontaneous emission rate.

A. Numerical test of the theory in predicting the complex resonance frequencies of the QNMs of the coupled-MNR system

The validity of the theory in Sec. III B in predicting the complex resonance frequencies of the QNMs of the coupled-MNR system will be tested in this subsection. The dispersive permittivity of gold is described by the Drude model (see Sec. II for model parameters) for numerical calculation. Then Eq. (24) can be used to calculate the complex resonance frequencies of the QNMs of the system.

As shown in Figs. 4(a)–4(c), we first discuss the coupling between two MNRs. The geometric parameters of each single MNR are the same as those in Fig. 1(a). For a single MNR, its spectral response in the spectral range of interest is dominated by the excitation of one QNM with the highest Q factor [with the complex resonance frequency $\tilde{\omega}_0/(2\pi c) = 0.8842 - 0.0477i$ μm^{-1}]. Therefore, only this QNM is considered for each MNR in our theoretical model. The coupled MNRs considered in Figs. 4(a) and 4(b) rigorously satisfy $\kappa_{(1,1),(2,1)} = \kappa_{(2,1),(1,1)}$, and the coupled MNRs considered in Fig. 4(c) approximately satisfy $\kappa_{(1,1),(2,1)} \approx \kappa_{(2,1),(1,1)}$. According to the theory in Sec. III D, one only needs to consider the two dominant QNMs with the highest Q factors supported by these MNR dimers, which are termed symmetric and antisymmetric modes, respectively [see Eq. (35) for the reason].

As shown in Figs. 4(a)–4(c), the complex resonance frequencies $\tilde{\omega}_r$ of the symmetric (blue curves) and antisymmetric (red curves) QNMs of the coupled system are plotted as a function of the distance d between the two MNRs. For large values of d, the model predictions (solid curves) agree well with the full-wave a-FMM results (circles), which confirms the validity of the theoretical model. With the decrease of d [e.g., $d < 60$ nm in Fig. 4(b)], the coupling effect between the two MNRs increases and the deviation between the model predictions and the full-wave data increases gradually. This deviation comes from the fact that only one QNM with the highest Q factor is considered for each MNR in the model, while all the other higher-order QNMs with lower Q factors are neglected.

As shown in Figs. 4(a2)–4(c2), for $d < 100$ nm the frequency shift $|\text{Re}(\Delta\tilde{\omega}_r)|$ (with $\Delta\tilde{\omega}_r = \tilde{\omega}_r - \tilde{\omega}_0$) of the real part of the complex resonance frequency $\tilde{\omega}_r$ (corresponding to the real resonance frequency of the coupled MNRs) increases with the decrease of d (the value of $\tilde{\omega}_0$ for a single MNR being shown by the horizontal black dotted lines). This phenomenon has been observed in experiments [94,95,98] and can be explained with our model as stated below. The expression of the frequency shift $|\text{Re}(\Delta\tilde{\omega}_r)| = |\text{Re}[\kappa(\omega)]|$ is given by the right-hand side of Eq. (47) in Sec. III D. In the expression, the overlap integral M_0 of the two QNMs considered in the model strongly depends on d since the integral is performed in the region of each MNR, the frequency $\omega = \tilde{\omega}_r$ is the slowly varying function of d, and F_0 is independent of d. For a system with low loss, the electric and magnetic fields of a QNM with a high Q factor approximately take real and purely imaginary values [i.e., $\tilde{\mathbf{E}}_m = \tilde{\mathbf{E}}_m^*$ and $\tilde{\mathbf{H}}_m = -\tilde{\mathbf{H}}_m^*$ as mentioned after Eq. (5)], respectively, which leads F_0 and M_0 to take real values. Thus, with the decrease of d and for small values of d, M_0 increases [see Eq. (42) for the definition of M_0] so that $|\text{Re}(\Delta\tilde{\omega}_r)|$ also increases.

Besides, Figs. 4(a2)–4(c2) show that for $d < 100$ nm, there are $|\text{Re}(\Delta\tilde{\omega}_+)| < |\text{Re}(\Delta\tilde{\omega}_-)|$ for $\text{Re}(\tilde{\omega}_+) > \text{Re}(\tilde{\omega}_-)$ (subscripts $+$ and $-$ corresponding to $r = 1$ and 2, respectively), and $|\text{Re}(\Delta\tilde{\omega}_+)| > |\text{Re}(\Delta\tilde{\omega}_-)|$ for $\text{Re}(\tilde{\omega}_+) < \text{Re}(\tilde{\omega}_-)$, which are consistent with the conclusion predicted by Eq. (47) in Sec. III D.

Next we will consider a more general system of a trimer antenna composed of three coupled MNRs, as sketched in Fig. 4(d1). The same as before, only the QNM with the highest Q factor is considered for each individual MNR in the theoretical model. This QNM will split into three QNMs with high Q factors due to the coupling. Their complex resonance frequencies plotted as functions of spacing d between MNRs are shown in Figs. 4(d2)–4(d3), where the curves with different colors correspond to different orders of QNMs. The results show that for large values of d($d > 60$ nm), the model predictions (pluses) are consistent well with the a-FMM numerical results (circles), which confirms the validity of the model. The error of the model increases gradually with the decrease of d, which is similar to the aforementioned case of the MNR dimer. In addition, the mode frequency shift $|\text{Re}(\Delta\tilde{\omega}_r)|$ increases as d decreases for $d < 100$ nm.

For the above examples, in the model only one QNM is considered for each MNR in the coupled-MNR system. To test the validity of our theory for the case that multiple QNMs are considered for each single MNR, we treat the trimer antenna shown in Fig. 4(d1) as composed of a dipole antenna [Fig. 4(a1)] and a single-wire antenna. According to the results in Fig. 4(a), the dipole antenna supports two dominant QNMs with high Q factors. They are considered in the model and are obtained with full-wave a-FMM numerical calculations (instead of with the QNM coupling theory). For the single-wire antenna, only one QNM with the highest Q factor is considered in the model as before. Figure 4(d) shows that the model predictions (solid curves) agree well with the a-FMM numerical results (circles), which confirms the validity of the model. Compared to the previous model that treats the trimer antenna as composed of three MNRs [pluses in Fig. 4(d)], the present model (solid curves) also shows good consistency for large values of the MNR spacing d($d > 60$ nm) .

But with the decrease of d, there appears difference between the predictions of the two models. The reason is that for the present model, the two dominant QNMs with the highest Q factors supported by the dipole antenna are obtained with full-wave numerical calculations, instead of considering the coupling between MNRs as done in the previous model. For large values of d, these two QNMs can be well reproduced through the coupling of the QNM with the highest Q factor supported by a single-wire antenna [as shown in Figs. 4(a2)–4(a3)], so that the predictions of the two models become consistent. While for small values of d, the accuracy of the reproduction is low, which results in the difference between the predictions of the two models.

For the above case that the trimer antenna is treated as composed of a dipole antenna and a single-wire antenna, we further consider the situation that the gap size of the dipole antenna is fixed to be a small value of $l = 10$ nm, so as to further test the validity of the theory. Figure 4(e) shows the complex resonance frequencies $\tilde{\omega}_r$ of the three dominant

QNMs with the highest Q factors supported by the trimer antenna, which are plotted as functions of the distance d between the dipole antenna and the single-wire antenna. It is seen that for large values of $d (d > 60$ nm), the predictions of the model are in good agreement with the results of the full-wave a-FMM. However, for this case, a larger error will occur for the model if the trimer antenna is treated as composed of three single-wire antennas, which is due to the very small distance l [as shown in Fig. 4(a)].

Figures 4(a)–4(d) show that with the increase of d, the complex resonance frequency $\tilde{\omega}_r$ of QNMs of the coupled-MNR system will tend to the complex resonance frequency $\tilde{\omega}_0$ of the single-wire antenna (as shown with the horizontal black dotted lines). Figure 4(e) shows that with the increase of d, $\tilde{\omega}_r$ of the trimer antenna will tend to the complex resonance frequencies of the QNMs of the single-wire antenna (horizontal blue dotted lines) and of the dipole antenna (horizontal red and green dotted lines). That is, the complex resonance frequency $\tilde{\omega}_r$ of the coupled-MNR system will tend to the complex resonance frequency $\tilde{\omega}_{p,m}$ of a single MNR as d increases. This phenomenon can be explained by our theory. Equation (17) indicates that if d is large enough, there is $\kappa_{(p,m),(q,n)}(\omega) \approx 0$ due to $\tilde{\mathbf{E}}_{q,n}(\mathbf{r}) \approx 0$ within the region V_p of the pth MNR. Thus, the matrix $\mathbf{K}$ in Eq. (23) becomes a diagonal matrix with the diagonal element $\delta_{p,q}\delta_{m,n}\tilde{\omega}_{p,m}$, and then the solution of Eq. (23) is simply $\omega = \tilde{\omega}_r = \tilde{\omega}_{p,m}$.

On the other hand, the QNM field diverges at infinity as stated in Sec. II. This causes that for $d \to \infty$, there is $\kappa_{(p,m),(q,n)}(\omega) \to \infty$ so that $\tilde{\omega}_r \to \tilde{\omega}_0$ does not hold any longer. As shown in Fig. S1 in Supplemental Material Sec. S5 [69], for the dipole antenna $\tilde{\omega}_r$ first approaches $\tilde{\omega}_0$ and then diverges with the increase of $d (d > 10 \mu$m), which confirms the above conclusion. This divergence problem of the QNM field is also pointed out in the QNM coupling theories in [51,52]. The reason for this phenomenon of divergence is that only one QNM with the highest Q factor is used in the model to expand the scattered field of a single MNR, while the other higher-order QNMs are all neglected, which leads to the inability to accurately reproduce the scattered field far away from the MNR. However, since the QNM of single MNR considered in the model has a high Q factor, the QNM field diverges only for the case that d is very large, for which generally the coupling effect between the scattered fields of MNRs can be neglected [i.e., $\kappa_{(p,m),(q,n)}(\omega) \approx 0$ as mentioned above]. Then the QNMs of the coupled-MNR system reduce to QNMs of each individual MNR in the system (diabolic-point degeneracy [99] will occur if the MNRs in the system are the same), so that it is unnecessary to solve the QNM coupling equations.

If some QNM of single MNR considered in the model has a low Q factor, its field diverges even for a small value of d. To solve this problem, one can first expand the scattered field inside the MNR (or on the surface of the MNR) with a high precision by using such QNM with a low Q factor, and then by using the Dyson's equation [100] (or the field equivalence principle [51]), one can obtain the scattered field far away from the MNR as the radiation field by an equivalent volume (or surface) current. Note that the above solution is valid only if the QNM of the coupled MNRs to be solved has a high Q factor. If not, the radiation field by the equivalent current,

which is at the complex eigenfrequency of the QNM of the coupled MNRs, still suffers from the divergence problem.

Besides providing some physical understanding, the theoretical model also has superior computational efficiency compared with full-wave numerical methods when changing the combination parameters (such as the MNRs' distance d in Fig. 4) of the MNRs in forming the coupled-MNR system, which is due to the analyticity of the model with respect to these combination parameters [see Eqs. (17) and (18)]. To illustrate the point, Table I provides a comparison of the computation times of different methods. One can see that for the data in Fig. 4(a), the computation time of the model (18.779 min) is much shorter than that of the full-wave a-FMM (8.781 h). For more details, $t_{\text{QNM}} = 29.271$ min and $t_{\text{QNM,0}} = 18.767$ min are the computation times of the a-FMM to solve the QNMs of the coupled MNRs and of the single MNR, respectively. The truncated harmonic orders in the transversal x and y directions for the a-FMM are $M_x = 30$ and $M_y = 30$, respectively (t_{QNM} and $t_{\text{QNM,0}}$ being proportional to $(M_x M_y)^3$ [101]). $t_{\text{K}} = 0.0663$ s and $t_{\text{C}} = 0.0118$ s are the computation times for calculating the matrix $\mathbf{K}$ of QNM coupling coefficients in Eq. (20) and for solving the eigenvalue problem (24) of the coupled MNRs, respectively. Note that t_{K} and t_{C} will increase with the number of QNMs considered in the model for the single MNR. The factor 2 is the number of QNMs solved for the dipole antenna. Besides the computation time, the consumed memory can be also saved with the model than with the full-wave numerical methods, since the former only requires a full-wave calculation of the QNMs of the single MNR (corresponding to $t_{\text{QNM,0}}$ in Table I), while the latter requires a full-wave calculation of the QNMs of the coupled MNRs (corresponding to t_{QNM}), which generally have a larger spatial extent than the single MNR.

The above numerical test is for systems with a homogeneous background medium, while the model can apply as well to the system with an inhomogeneous background medium. To demonstrate the point, we provide an additional numerical example in Supplemental Material Sec. S6 [69], for which the MNR is an aperture carved in a layered background medium.

B. Numerical test of the theory in predicting the fields of the QNMs of the coupled-MNR system

In this subsection, we will reproduce the fields of the QNMs of the five coupled MNRs in Fig. 4 with the theoretical model. Figure 5 shows the electric-field distributions of the symmetric and antisymmetric QNMs of the dipole antenna [as sketched in Fig. 4(a1)] with $d = 200$, 100, and 60 nm. Here we show the E_z component which is the dominant electric-field component in the nanogap. For the symmetric QNM, the E_z components at the two antenna arms are in-phase, which thus forms a constructive interference and leads to an enhancement of the electric field in the nanogap. However, the antisymmetric QNM is of the opposite case with a weak electric field in the nanogap. The predictions of the model (the second row in Fig. 5) are in good agreement with the full-wave a-FMM calculation results (the first row) except for the deviation in the nanogap for $d = 60$ nm, which is consistent with the results in Figs. 4(a2) and 4(a3).

TABLE I. Computation times of different methods. The calculation is performed on a laptop computer with 4-core CPUs of 1.6 GHz and a memory of 12 GB. n is the number of data points.

Figure	Method	Computation time
Fig. 4(a) ($n = 9 \times 2 = 18$)	a-FMM for solving QNM	nt_{QNM} $= 18 \times 29.271$ min $= 8.781$ h
	model	$t_{QNM,0} + (n/2)(t_K + t_C)$ $= 18.767$ min $+ (18/2) \times (0.0663$ s $+ 0.0118$ s$) = 18.779$ min
Fig. 6(a) ($n = 50 \times 3 = 150$)	a-FMM	$nt_{a-FMM} = 150 \times 4.471$ min $= 11.177$ h
	QNM expansion	$2 \times 3 \times t_{QNM} = 2 \times 3 \times 29.271$ min $= 2.927$ h
	model 1	$t_{QNM,0} + 3t_K + n(t_b + t'_C)$ $= 18.767$ min $+ 3 \times 0.0663$ s $+150 \times (0.0341$ s $+ 0.002\,51$ s$)$ $= 18.862$ min
	model 2	$t_{QNM,0} + 3(t_K + 2t_b + t_C)$ $= 18.767$ min $+ 3 \times (0.0663$ s $+ 2$ $\times 0.0341$ s $+ 0.0118$ s$)$ $= 18.774$ min

For the other four antennas calculated in Figs. 4(b)–4(e), Figs. S6–S9 in Supplemental Material Sec. S7 [69] show that the fields of their supported QNMs can be well reproduced with the model. For smaller values of spacing d between MNRs ($d = 60$ nm for instance), the model exhibits a larger error in reproducing the QNM field in the nanogap, which reflects a larger contribution of higher-order QNMs that are neglected in the model.

In Supplemental Material Sec. S8 [69], we provide a further numerical test of the theory in predicting the QNM expansion coefficients, i.e., the nontrivial solution $\mathbf{a} = \tilde{\mathbf{a}}_r$ of Eq. (22).
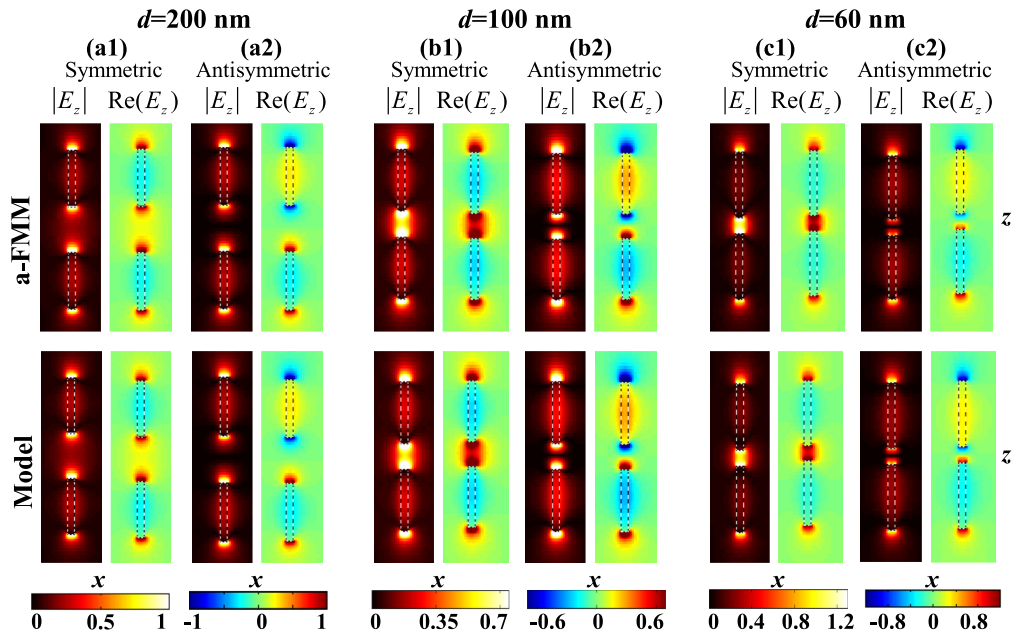

FIG. 5. Electric-field distributions of symmetric (a1)–(c1) and antisymmetric (a2)–(c2) QNMs [with the complex resonance frequencies shown in Fig. 4(a)] supported by the dipole antenna [as sketched in Fig. 4(a1)] for $d = 200$ (a), 100 (b), and 60 nm (c). The distributions of $|E_z|$ (left) and $\mathrm{Re}(E_z)$ (right) on $y = 0$ [the coordinate being shown in Fig. 4(a1)] are provided. The QNMs satisfy the normalization $E_z(0, 0, 0) = 1$, with the coordinate origin shown in Fig. 4(a1) as O. The first and second rows show the full-wave a-FMM results and the QNM coupling model predictions, respectively.
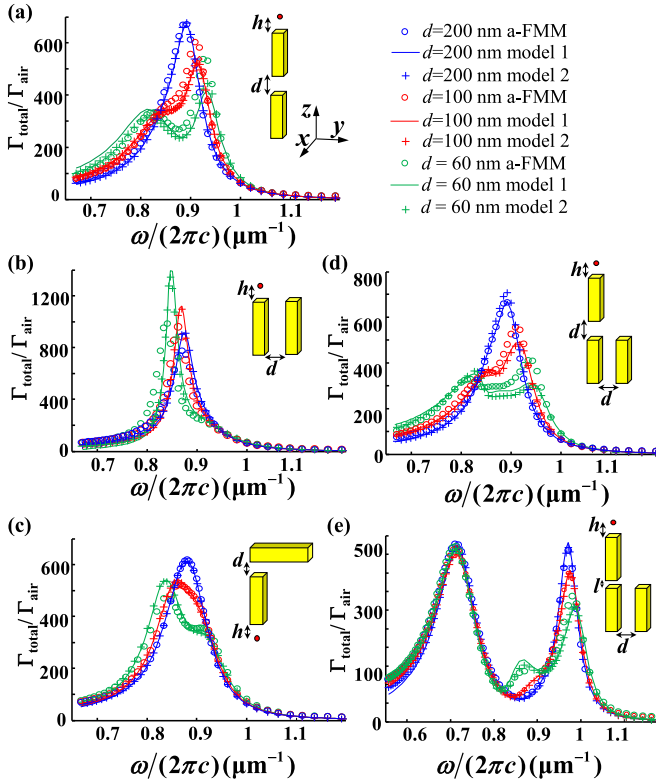
FIG. 6. Spectra of the normalized total emission rate $\Gamma_{\text{total}}/\Gamma_{\text{air}}$ obtained with the QNM coupling model (solid curves and pluses, corresponding to the theories proposed in Secs. III A and C, respectively) and with the full-wave a-FMM (circles). The results in (a)–(e) are given in turn for the dipole antenna, parallel-nanowires antenna, vertical-nanowires antenna, and trimer antenna as sketched in Figs. 4(a1)–4(e1), respectively. $l = 10$ nm is fixed in (e). The red point in the insets represents a z-polarized point current source located at $h = 10$ nm away from the antenna end face (its projection on the antenna end face being at the center of the end face). The blue, red, and green curves correspond to $d = 200$, 100, and 60 nm, respectively.

C. Numerical test of the theory in predicting the source-excited field of the coupled-MNR system

In this subsection, we will test the validity of the QNM coupling theory proposed in Secs. III A and III C in the presence of excitation sources. For the five coupled MNRs calculated in Fig. 4, the spectrum of the total emission rate Γ_{total} of a point current source near the coupled MNRs are calculated, as shown in Fig. 6, which provides the results for different spacings $d = 200$ (blue curves), 100 (red), and 60 nm (green) between MNRs. Here the point current source is used to represent a molecule or quantum dot with fluorescence emission. In order to simultaneously excite the QNMs with different symmetries, we appropriately locate the point source to break the symmetry of the system, as shown with the red dot in Fig. 6.

In the theoretical model, we only consider one dominant QNM with the highest Q factor supported by each nanowire in the coupled MNRs. In Fig. 6, model 1 and model 2 correspond to the theories in Secs. III A and III C, respectively. Note that the calculation of the background field with different

real excitation frequencies ω [i.e., $\mathbf{E}^b(\mathbf{r}, \omega)$ in Eq. (18)] are required in model 1, while only the background field with $\omega = \tilde{\omega}_r$ is required in model 2 [see Eq. (31)]. Since only a few dominant QNMs with the highest Q factors (i.e., QNMs calculated in Fig. 4) supported by the coupled MNRs are considered in model 2, the results of model 1 and model 2 will not be exactly the same.

Figure 6 shows that the predictions of model 1 (solid curves) and model 2 (pluses) are both consistent with the full-wave a-FMM results (circles), thus confirming the validity of the theories proposed in Secs. III A and III C. In addition, the consistent predictions of the two models indicate that the contribution of the QNMs neglected in model 2 is small.

Figure 4 shows that the difference between the resonance frequencies of the splitting QNMs of the coupled MNRs increases as the spacing d between MNRs decreases ($d = 100$ and 60 nm for instance) due to the stronger coupling. On the other hand, these QNMs will be excited by the point current source, resulting in two or more resonance peaks (the peak frequencies being the real part of the complex eigenfrequencies) in the spectra of Γ_{total} as shown in Fig. 6. Therefore, a broader band of enhancement of the spontaneous emission rate Γ_{total} can be achieved for a smaller d. For example, for the trimer antenna with $d = 60$ nm as shown in Fig. 6(d), the wavelength range of $\Gamma_{\text{total}}/\Gamma_{\text{air}} > 200$ is 320 nm, while the corresponding wavelength range for the single-wire antenna is only 190 nm. Since the fluorescence wavelength of molecules or quantum dots usually covers a range from tens to hundreds of nanometers, the broadband enhancement of the spontaneous emission is of great significance for high-sensitivity sensing of fluorescent molecules [102], high-brightness LED light source [103], nanolaser [104], and so on. In addition, the coupled MNRs with multiple resonance frequencies can be used to enhance both the excitation rate at the fluorescence excitation frequency and the spontaneous emission rate at the fluorescence emission frequency, so as to further enhance the fluorescence intensity [12,105].

Concerning the computational efficiency, Table I shows that for the data in Fig. 6(a), the computation times of models 1 and 2 (18.862 and 18.774 min, respectively) are much shorter than that of the QNM expansion method (2.927 h) for which the QNMs of the coupled MNRs are solved with the full-wave a-FMM. This comparison shows the superior computational efficiency of the theoretical models when changing the combination parameters (such as the MNRs' distance d in Fig. 6) of the MNRs in forming the coupled-MNR system. For more details, the definitions and values of t_{QNM}, $t_{\text{QNM,0}}$, t_K, and t_C are the same as those for the data of Fig. 4(a). $t_b = 0.0341$ s and and $t'_C = 0.00251$ s are the computation times for calculating the vector $\mathbf{b}$ of the QNM excitation coefficients in the coupling equation (20) and for solving this equation, respectively. The factors 3 and 2 are the number of d's values and the number of QNMs solved for the dipole antenna, respectively. As expected, Table I also shows that the computation time of the QNM expansion method (2.927 h) is much shorter than that of the full-wave a-FMM (11.177 h), which is due to the frequency analyticity of the former as already mentioned close to the end of Sec. II and for the data in Fig. 1(c).

V. CONCLUSION

For the coupled-MNR system composed of dispersive and lossy MNRs, we analyze the coupling mechanism based on the Green's-function formalism and show that the coupling between MNRs originates from a mutual excitation of the scattered field of each MNR. Based on this point, a QNM coupling theory is built up with the use of the QNM expansion formalism. Compared with the existing coupled mode theories, our theory can be generally applied to a coupling system with energy loss and an arbitrary frequency dispersion of permittivity. Besides, all the parameters in the theory (such as the QNM coupling coefficients and excitation coefficients) are calculated based on the first principle of Maxwell's equations, so that there is no need to fit the full-wave numerical calculation results or experimental data, which ensures a solid electromagnetic foundation of the theory.

With the theory, analytical expressions of the complex resonance frequency and field distribution of the QNMs of the coupled-MNR system in terms of the QNMs of each single MNR are obtained. In the presence of excitation source, analytical expression of the scattered field of the coupled-MNR system in terms of the excitation source and the QNMs of each single MNR is obtained, and the scattered field of the coupled-MNR system expanded upon the basis of the QNMs of the coupled-MNR system is further obtained via the Mittag-Leffler theorem.

For several representative coupled MNRs, their supported QNMs in the absence of excitation source and their electromagnetic response in the presence of excitation source are calculated. The validity of the proposed theoretical model is confirmed through a comparison between the model predictions and the full-wave a-FMM numerical results. Based on the theory, the properties of the QNMs and the electromagnetic response of the coupled MNRs are explained, and

coupled-MNR systems with a broadband enhancement of the spontaneous emission rate are designed. The proposed theory can provide intuitive and quantitative physical pictures for understanding and designing the optical properties of coupled MNRs for many applications, such as the enhancement of spontaneous emission rate [102–104], plasmon ruler [19,20], exceptional point [25–27], and plasmon-induced transparency [23,24]. Besides, the theory also shows superior computational efficiency compared with full-wave numerical methods when changing the combination (such as the number and relative position) of the MNRs in forming the coupled-MNR system.

The range of validity and the limitation of the theoretical model are summarized below. With the decrease of the distance between MNRs, the accuracy of the model will decrease if only the dominant QNMs of single MNRs with high Q factors are considered in the model. The accuracy can be improved by considering more QNMs of single MNRs with low Q factors, but at the expense of increased computational amount and less physical intuitiveness. On the other hand, the far-field divergence of QNMs of single MNRs at complex eigenfrequencies will appear with the increase of the distance between MNRs. The present theory is limited to isotropic and nonmagnetic medium, but can be readily extended to the more general anisotropic or bianisotropic medium [62], or even nonreciprocal medium.

ACKNOWLEDGMENTS

[1] D. K. Gramotnev and S. I. Bozhevolnyi, Nat. Photon. **4**, 83 (2010).

[2] S. Kuhn, U. Hakanson, L. Rogobete, and V. Sandoghdar, Phys. Rev. Lett. **97**, 017402 (2006).

[3] J. N. Anker, W. P. Hall, O. Lyandres, N. C. Shah, J. Zhao, and R. P. Van Duyne, Nat. Mater. **7**, 442 (2008).

[4] S. Y. Ding, E. M. You, Z. Q. Tian, and M. Moskovits, Chem. Soc. Rev. **46**, 4042 (2017).

[5] L. Tang, S. E. Kocabas, S. Latif, A. K. Okyay, D.-S. Ly-Gagnon, K. C. Saraswat, and D. A. B. Miller, Nat. Photon. **2**, 226 (2008).

[6] Y. De Wilde, F. Formanek, R. Carminati, B. Gralak, P.-A. Lemoine, K. Joulain, J.-P. Mulet, Y. Chen, and J.-J. Greffet, Nature (London) **444**, 740 (2006).

[7] S. Kim, N. Yu, X. Ma, Y. Zhu, Q. Liu, M. Liu, and R. Yan, Nat. Photon. **13**, 636 (2019).

[8] R. F. Oulton, V. J. Sorger, T. Zentgraf, R.-M. Ma, C. Gladden, L. Dai, G. Bartal, and X. Zhang, Nature (London) **461**, 629 (2009).

[9] M. A. Noginov G. Zhu, A. M. Belgrave, R. Bakker, V. M. Shalaev, E. E. Narimanov, S. Stout, E. Herz, T. Suteewong, and U. Wiesner, Nature (London) **460**, 1110 (2009).

[10] L. Novotny and N. van Hulst, Nat. Photon. **5**, 83 (2011).

[11] H. Aouani, M. Rahmani, H. Šípová, V. Torres, K. Hegnerová, M. Beruete, J. Homola, M. Hong, M. Navarro-Cía, and S. A. Maier, J. Phys. Chem. C **117**, 18620 (2013).

[12] G. Lu, J. Liu, T. Zhang, H. Shen, P. Perriat, M. Martini, O. Tillement, Y. Gu, Y. He, Y. Wang *et al.*, Nanoscale **5**, 6545 (2013).

[13] A. G. Curto, G. Volpe, T. H. Taminiau, M. P. Kreuzer, R. Quidant, and N. F. van Hulst, Science **329**, 930 (2010).

[14] V. V. Thacker, L. O. Herrmann, D. O. Sigle, T. Zhang, T. Liedl, J. J. Baumberg, and U. F. Keyser, Nat. Commun. **5**, 3448 (2014).

[15] A. Kinkhabwala, Z. Yu, S. Fan, Y. Avlasevich, K. Müllen, and W. E. Moerner, Nat. Photon. **3**, 654 (2009).

[16] D. Punj, M. Mivelle, S. B. Moparthi, T. S. van Zanten, H. Rigneault, N. F. van Hulst, M. F. Garcia-Parajo, and J. Wenger, Nat. Nanotechnol. **8**, 512 (2013).

[17] H. Harutyunyan, G. Volpe, R. Quidant, and L. Novotny, Phys. Rev. Lett. **108**, 217403 (2012).

[18] B. Metzger, M. Hentschel, T. Schumacher, M. Lippitz, X. Ye, C. B. Murray, B. Knabe, K. Buse, and H. Giessen, Nano Lett. **14**, 2867 (2014).

[19] B. M. Reinhard, M. Siu, H. Agarwal, A. P. Alivisatos, and J. Liphardt, Nano Lett. **5**, 2246 (2005).

[20] N. Liu, M. Hentschel, T. Weiss, A. P. Alivisatos, and H. Giessen, Science **332**, 1407 (2011).

[21] J. Aaron, K. Travis, N. Harrison, and K. Sokolov, Nano Lett. **9**, 3612 (2009).

[22] M. Liu, Q. Li, L. Liang, J. Li, K. Wang, J. Li, M. Lv, N. Chen, H. Song, J. Lee *et al.*, Nat. Commun. **8**, 15646 (2017).

[23] S. Zhang, D. A. Genov, Y. Wang, M. Liu, and X. Zhang, Phys. Rev. Lett. **101**, 047401 (2008).

[24] X. Niu, X. Hu, Q. Yan, J. Zhu, H. Cheng, Y. Huang, C. Lu, Y. Fu, and Q. Gong, Nanophotonics **8**, 1125 (2019).

[25] A. Kodigala, T. Lepetit, and B. Kanté, Phys. Rev. B **94**, 201103(R) (2016).

[26] S. K. Ozdemir, S. Rotter, F. Nori, and L. Yang, Nat. Mater. **18**, 783 (2019).

[27] J.-H. Park, A. Ndao, W. Cai, L. Hsu, A. Kodigala, T. Lepetit, Y.-H. Lo, and B. Kanté, Nat. Phys. **16**, 462 (2020).

[28] H. A. Haus and W. Huang, Proc. IEEE **79**, 1505 (1991).

[29] B. Xi, H. Xu, S. Xiao, and L. Zhou, Phys. Rev. B **83**, 165115 (2011).

[30] N. A. Gippius, T. Weiss, S. G. Tikhodeev, and H. Giessen, Opt. Express **18**, 7569 (2010).

[31] F. Alpeggiani, N. Parappurath, E. Verhagen, and L. Kuipers, Phys. Rev. X **7**, 021035 (2017).

[32] E. Prodan, Science **302**, 419 (2003).

[33] P. Nordlander, C. Oubre, E. Prodan, K. Li, and M. I. Stockman, Nano Lett. **4**, 899 (2004).

[34] E. Prodan and P. Nordlander, J. Chem. Phys. **120**, 5444 (2004).

[35] T. J. Davis, K. C. Vernon, and D. E. Gómez, Phys. Rev. B **79**, 155423 (2009).

[36] B. Xi, M. Qiu, S. Xiao, H. Xu, and L. Zhou, Phys. Rev. B **89**, 035110 (2014).

[37] M. Qiu, S. Xiao, Q. He, S. Sun, and L. Zhou, Opt. Lett. **40**, 272 (2015).

[38] H. A. Haus, *Waves and Fields in Optoelectronics* (Central Book Co., Taipei, 1985).

[39] S. Wonjoo, W. Zheng, and S. Fan, IEEE J. Quantum Electron. **40**, 1511 (2004).

[40] A. Artar, A. A. Yanik, and H. Altug, Nano Lett. **11**, 3694 (2011).

[41] R. Yang, Q. Fu, Y. Fan, W. Cai, K. Qiu, W. Zhang, and F. Zhang, Photon. Res. **7**, 955 (2019).

[42] G. Garcia-calderón and R. Peierls, Nucl. Phys. A **265**, 443 (1976).

[43] S. C. Hill and R. E. Benner, *Morphology-dependent Resonances* (World Scientific, Singapore, 1988).

[44] B. R. Johnson, J. Opt. Soc. Am. A **10**, 343 (1993).

[45] P. T. Leung, S. Y. Liu, and K. Young, Phys. Rev. A **49**, 3057 (1994).

[46] E. S. C. Ching, P. T. Leung, A. Maassen van denBrink, W. M. Suen, S. S. Tong, and K. Young, Rev. Mod. Phys. **70**, 1545 (1998).

[47] P. Lalanne, W. Yan, K. Vynck, C. Sauvan, and J.-P. Hugonin, Laser Photon. Rev. **12**, 1700113 (2018).

[48] H. Jia, F. Yang, Y. Zhong, and H. Liu, Photon. Res. **4**, 293 (2016).

[49] P. T. Kristensen, J. R. de Lasson, M. Heuck, N. Gregersen, and J. Mørk, J. Lightwave Technol. **35**, 4247 (2017).

[50] B. Vial and Y. Hao, J. Opt. **18**, 115004 (2016).

[51] P. T. Kristensen, K. Herrmann, F. Intravaia, and K. Busch, arXiv:1910.05412.

[52] K. Cognee, Ph.D. dissertation, Univ. of Bordeaux, Bordeaux, 2020.

[53] S. A. Maier, Opt. Quantum Electron. **38**, 257 (2006).

[54] A. F. Koenderink, Opt. Lett. **35**, 4208 (2010).

[55] A. Raman and S. Fan, Phys. Rev. Lett. **104**, 087401 (2010).

[56] M. Garcia-Vergara, G. Demesy, and F. Zolla, Opt. Lett. **42**, 1145 (2017).

[57] E. A. Muljarov and W. Langbein, Phys. Rev. B **93**, 075417 (2016).

[58] D. Bergman, Phys. Rev. B **19**, 2359 (1979).

[59] M. Pascale, G. Miano, R. Tricarico, and C. Forestiere, Sci. Rep. **9**, 14524 (2019).

[60] G. Rosolen, B. Maes, P. Y. Chen, and Y. Sivan, Phys. Rev. B **101**, 155401 (2020).

[61] P. Y. Chen, D. J. Bergman, and Y. Sivan, Phys. Rev. Appl. **11**, 044018 (2019).

[62] E. A. Muljarov and T. Weiss, Opt. Lett. **43**, 1978 (2018).

[63] F. Zolla, A. Nicolet, and G. Demesy, Opt. Lett. **43**, 5813 (2018).

[64] C. Sauvan, J. P. Hugonin, I. S. Maksymov, and P. Lalanne, Phys. Rev. Lett. **110**, 237401 (2013).

[65] W. Yan, R. Faggiani, and P. Lalanne, Phys. Rev. B **97**, 205422 (2018).

[66] A. Gras, P. Lalanne, and M. Duruflé, J. Opt. Soc. Am. A **37**, 1219 (2020).

[67] C. Gigli, T. Wu, G. Marino, A. Borne, G. Leo, and P. Lalanne, ACS Photon. **7**, 1197 (2020).

[68] S. Franke, S. Hughes, M. K. Dezfouli, P. T. Kristensen, K. Busch, A. Knorr, and M. Richter, Phys. Rev. Lett. **122**, 213901 (2019).

[69] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.102.045430 for the numerical calculation method of the pseudoenergy F_m (Sec. S1), the detailed derivation of Eq. (24) (Sec. S2), a logic crosscheck on the validity of the QNM coupling theory (Sec. S3), more details concerning the relation between our theory and other theories (Sec. S4), and additional numerical test of the QNM coupling theory (Secs. S5–S8).

[70] K. Zhang and D. Li, *Electromagnetic Theory for Microwaves and Optoelectronics* (Springer, Berlin, 2008).

[71] Q. Bai, M. Perrin, C. Sauvan, J. P. Hugonin, and P. Lalanne, Opt. Express **21**, 27371 (2013).

[72] P. Lalanne *et al.*, J. Opt. Soc. Am. A **36**, 686 (2019).

[73] B. Gralak and A. Tip, J. Math. Phys. **51**, 052902 (2010).

[74] Y. Brûlé, B. Gralak, and G. Demésy, J. Opt. Soc. Am. B **33**, 691 (2016).

[75] Y. Li, H. Liu, H. Jia, F. Bo, G. Zhang, and J. Xu, J. Opt. Soc. Am. A **31**, 2459 (2014).

[76] F. Binkowski, L. Zschiedrich, and S. Burger, J. Comput. Phys. **419**, 109678 (2020).

[77] T. Weiss, N. A. Gippius, S. G. Tikhodeev, G. Granet, and H. Giessen, J. Opt. Soc. Am. A **28**, 238 (2011).

[78] J. R. de Lasson, P. T. Kristensen, J. Mork, and N. Gregersen, J. Opt. Soc. Am. A **31**, 2142 (2014).

[79] E. D. Palik, *Handbook of Optical Constants of Solids Part II* (Academic, Boston, 1991).

[80] P. Muhlschlegel, H. J. Eisler, O. J. Martin, B. Hecht, and D. W. Pohl, Science **308**, 1607 (2005).

[81] L. Novotny and B. Hecht, *Principles of Nano-Optics* (Cambridge University Press, Cambridge, UK, 2006).

[82] E. M. Purcell, H. C. Torrey, and R. V. Pound, Phys. Rev. **69**, 674 (1946).

[83] M. Pelton, Nat. Photon. **9**, 427 (2015).

[84] J. P. Hugonin and P. Lalanne, J. Opt. Soc. Am. A **22**, 1844 (2005).

[85] The calculation is performed with an in-house software: H. Liu, *DIF CODE for Modeling Light Diffraction in Nanostructures* (Nankai University, Tianjin, 2010).

[86] A. B. Evlyukhin, G. Brucoli, L. Martín-Moreno, S. I. Bozhevolnyi, and F. J. García-Vidal, Phys. Rev. B **76**, 075426 (2007).

[87] F. Tisseur and K. Meerbergen, SIAM Rev. **43**, 235 (2001).

[88] J.-P. Dedieu and F. Tisseur, Linear Alg. Appl. **358**, 71 (2003).

[89] G. B. Arfken, H. J. Weber, and F. E. Harris, in *Mathematical Methods for Physicists*, 6th ed. (Academic, Boston, 2005).

[90] E. A. Muljarov and W. Langbein, Phys. Rev. B **94**, 235438 (2016).

[91] R. Baker and K. Kuttler, *Linear Algebra with Applications* (World Scientific, Singapore, 2014).

[92] M. Ringler, A. Schwemer, M. Wunderlich, A. Nichtl, K. Kurzinger, T. A. Klar, and J. Feldmann, Phys. Rev. Lett. **100**, 203002 (2008).

[93] E. E. Tyrtyshnikov, *A Brief Introduction to Numerical Analysis* (Springer, New York, 1997).

[94] W. Rechberger, A. Hohenau, A. Leitner, J. R. Krenn, B. Lamprecht, and F. R. Aussenegg, Opt. Commun. **220**, 137 (2003).

[95] A. M. Funston, C. Novo, T. J. Davis, and P. Mulvaney, Nano Lett. **9**, 1651 (2009).

[96] G. Lu, Y. Wang, R. Y. Chou, H. Shen, Y. He, Y. Cheng, and Q. Gong, Laser Photon. Rev. **9**, 530 (2015).

[97] B. Cohn, B. Engelman, A. Goldner, and L. Chuntonov, J. Phys. Chem. Lett. **9**, 4596 (2018).

[98] K. H. Su, Q. H. Wei, X. Zhang, J. J. Mock, D. R. Smith, and S. Schultz, Nano Lett. **3**, 1087 (2003).

[99] M. V. Berry and M. Wilkinson, Proc. R. Soc. Lond. A **392**, 15 (1984).

[100] R.-C. Ge, P. T. Kristensen, J. F. Young, and S. Hughes, New J. Phys. **16**, 113048 (2014).

[101] B. Bai and L. Li, J. Opt. Soc. Am. A **21**, 1886 (2004).

[102] A. F. Koenderink, ACS Photon. **4**, 710 (2017).

[103] K. L. Tsakmakidis, R. W. Boyd, E. Yablonovitch, and X. Zhang, Opt. Express **24**, 17916 (2016).

[104] Z. Wang, X. Meng, A. V. Kildishev, A. Boltasseva, and V. M. Shalaev, Laser Photon. Rev. **11**, 1700212 (2017).

[105] K. Dopf, C. Moosmann, S. W. Kettlitz, P. M. Schwab, K. Ilin, M. Siegel, U. Lemmer, and H.-J. Eisler, ACS Photon. **2**, 1644 (2015).