Inverse Design of Focused Vector Beams for Mode Excitation in Optical Nanoantennas

Xiaorun Zang[*], Ari T. Friberg, Tero Setälä, and Jari Turunen

Institute of Photonics, University of Eastern Finland, P.O. Box 111, Joensuu FI-80101, Finland

We propose a free-space, inverse design of a nanostructure's effective mode-matching fields via a backward propagation of tightly focused vector beams to the pupil plane of an aplanatic system of high numerical aperture. First, we study the nanostructure's eigenmodes without considering any excitation fields and then extract the modal near fields in the focal plane. Each modal field is then taken as the desired focal field, the band-limited waves of which are backward propagated to the pupil plane via a reversal of the Richards-Wolf vector diffraction formula. The pupil fields can be designed to be genuinely paraxial by associating the longitudinal electric or magnetic field component with the radial one on the reference sphere. The inversely designed pupil field in turn is propagated forwardly into the focal region to generate the designed focal field, whose distribution over the nanostructure's surface is used to evaluate the overlap between the designed focal field and the modal fields, i.e., the modal expansion coefficients. Studies for a silicon nanodisk monomer, dimer, and tetramer demonstrate the ability of our inverse approach to design the necessary tightly focused vector field that can effectively and exclusively match a certain eigenmode of interest. Compared with the forward beam-shaping method, the inverse design approach tends to yield quantitatively more precise mode-matching field profiles. This work can have a significant impact on optical applications that rely on controllable and tunable mode excitation and light scattering.

I. INTRODUCTION

A strong local field in plasmonic or high-index dielectric nanostructures is beneficial for boosting a variety of optical effects, via an enhanced interaction between the strong light field and nanostructures. In contrast to plasmonic systems, high-index dielectric nanoparticles, such as silicon nanodisks [1–4], do not suffer from strong optical losses but still sustain deep subwavelength field confinements [5]. Moreover, all-dielectric nanostructures have been demonstrated to be versatile semiconductor-compatible nanophotonic devices for directional scattering [1], near-field mapping of optical modes [6], tailoring second-harmonic [7–9] and third-harmonic [10] generations, anapole-enhanced Raman scattering [3], etc. With recent advances, all-dielectric nanostructures form promising basic building blocks that complement or even replace the plasmonic counterparts in modern nanophotonics [5].

At certain frequencies the local field enhancement achieves maximum values (i.e., peaks), which is a phenomenon described as optical resonance [5,11]. It is known

that the resonance response can be tuned by varying the size, shape, and dielectric environment of the nanostructures [12]. The resonance is commonly analyzed with respect to a certain incident light field, such as a plane wave used in previous studies [1,3,13], which, however, may not probe all resonances [14]. Recently, it has been shown that vector beams [15,16] with spatially varying polarization distributions can reveal otherwise hidden resonances, such as the dark resonance detected by the radially or azimuthally polarized cylindrical vector beams [17–19], or the hybrid resonances by higher-order vector beams possessing helical phase distributions [20,21].

The optical resonance response in nanostructures is associated with the excitation of one or more modes [14] that, strictly speaking, are independent of any external excitation field and are hence intrinsically determined by the optical properties of the nanostructure and surrounding media. The resonance frequencies are generally isolated poles in the complex frequency plane, a characteristic of quasinormal modes [22]. At a fixed real-valued frequency, however, a set of discrete eigenmodes [14,23] can be defined. Modes in nanostructures have been investigated by various approaches, including the electrostatic method [24–26], and several rigorous full-wave analyses based on the Fourier modal method [27–29], transfer matrix method

---

[30], finite element method [22,31], Green's tensor method [32], boundary element method (BEM) based on the surface integral equations (SIEs) [14,23,33,34], etc. Among them, BEM proved to be efficient, particularly in modeling the interaction between the focused vector field and nanostructures [21,35,36], because the numerical discretization as well as the corresponding focal fields evaluation are reduced to two dimensions for a general three-dimensional (3D) problem.

The excitation efficiency of an eigenmode scales with the overlap between the modal and the applied incident field profiles, as well as the corresponding eigenvalue [14]. An eigenmode's near field in optical nanoantennas is vectorial and localized such that it may vary considerably within a subwavelength scale, which is a fact that poses a challenge to the choice of an excitation field with an effective mode-matching profile. A free-space approach to obtain such a mode-matching vector field is to shape a focal field that is to match with the modal field, by focusing a paraxial vector beam in an aplanatic system with high numerical aperture (NA). A forward design approach has been shown to yield spatially varying field distributions for efficient hybrid mode excitations [21], but a forwardly designed focal field is limited to a qualitative mode matching due to energy exchanges between the radial and azimuthal field components in the focusing process, and then a deviation of its field profile from the designed pupil field's. Alternatively, an inverse design is versatile in shaping focal fields that potentially bring more refined and effective mode-matching fields. The desired focal field can be made to match directly the eigenmode's near field in a certain plane in the inverse design approach. However, in the forward design method we can only match the incident pupil field with an eigenmode's near field. The desired incident pupil field changes its distribution during the tight focusing process due to the exchange of the radial and azimuthal field components and the designed focal field will not be the same.

In previous work of the inverse problem in a high-NA aplanatic system, interest has mainly focused on the focal field total intensity or a single component from which the complex pupil field is retrieved. For instance, the depth of focus can be optimized after determining the abberation and amplitude functions in a high-NA imaging system [37], or even retrieving the birefringence from four focal field intensity distributions generated by pupil fields of different polarization states [38]. A focal field with a null longitudinal component is generated by azimuthally polarized light [39]. Needle-, tube-, and bubble-shaped focal field intensity distributions [40], a perfect polarization vortex focal field [41], as well as a focal field of arbitrary homogeneous polarization [42] are also investigated. However, a mode-matching focal field generally involves more than one or all field components to be effectively matched with both the amplitude and phase distributions of the modal field, which has not been considered in previous work.

In this work, we present an inverse design of the desired focal field for effective mode excitation in optical nanoantennas by taking into account all modal field components and the inherent dependencies of all electric and magnetic field components in the designed beamlike pupil, reference, and focal fields. Independent of any excitation field, the eigenmodes are fully determined by the nanostructure's optical properties themselves and thus they are studied via a BEM mode solver implemented through the Müller formulation of the SIEs [43]. The modal fields in the focal plane are evaluated via the Stratton-Chu surface integral [44], and the extracted focal fields are propagated back to the pupil plane by reversing the tight focusing process. The best possible incident paraxial beams before focusing are then obtained from solving an inverse problem governed by the Richards-Wolf vector diffraction formula.

This work is organized as follows. In Sec. II, we study the eigenmodes in nanostructures, where we review the SIEs, describe the eigenvalue problem based on the Müller formulation, and analyze the eigenmodes in a silicon nanodisk monomer, dimer, and tetramer. In Sec. III, we describe the forward and backward tight focusing processes that are governed by the Richards-Wolf vector diffraction formula. The degrees of freedom and strategies in our inverse design are then discussed through a systematic study of the mode-matching field design in a silicon nanodisk monomer, dimer, and tetramer. In Sec. IV, we analyze the quality of the designed focal field in terms of modal expansion coefficients, i.e., the overlaps between the designed focal field and all the first twelve eigenmodes considered in each nanodisk oligomer. Finally, conclusions and perspectives are given in Sec. V.

II. EIGENMODES IN NANOSTRUCTURES

A. Surface integral equations

We first describe the eigenmode problem in optical nanostructures using the SIE formulation. Without loss of generality, the SIE formulation is reviewed for the case of a single arbitrarily shaped 3D nanostructure, since it is readily extended to more general cases of nano-object ensembles. As shown in Fig. 1, the nano-object occupying domain Ω_1 contains a homogeneous and isotropic medium with permittivity ϵ_1 and permeability μ_1. The free space in domain Ω_0 has optical properties ϵ_0 and μ_0. The monochromatic electric and magnetic fields of harmonic time dependence $e^{-i\omega t}$ in the scatterer and free space are denoted by $\{\mathbf{E}_1, \mathbf{H}_1\}$ and $\{\mathbf{E}_0, \mathbf{H}_0\}$, respectively, where ω is the angular frequency.

Applying the field equivalence principle [44–48] to the scatterer, we are interested in the fields $\{\mathbf{E}_1, \mathbf{H}_1\}$ in domain Ω_1 and thus the fields elsewhere are not of interest. In this regard, the fields $\{\mathbf{E}_0, \mathbf{H}_0\}$ in domain Ω_0 can take any
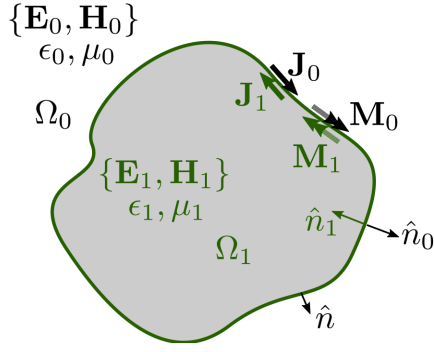
FIG. 1. An arbitrarily shaped 3D scatterer in free space. The scatterer that occupies domain Ω_1 has permittivity ϵ_1 and permeability μ_1, whereas domain Ω_0 is free space with optical properties ϵ_0 and μ_0. The electric and magnetic fields are $\{\mathbf{E}_0, \mathbf{H}_0\}$ in free space and $\{\mathbf{E}_1, \mathbf{H}_1\}$ in the scatterer. Two domains' surfaces are correspondingly denoted by $\partial\Omega_0$ and $\partial\Omega_1$, and their directions are determined by normal vectors $\hat{n}_0 = \hat{n}$ and $\hat{n}_1 = -\hat{n}$. Equivalent electric and magnetic surface current densities $\{\mathbf{J}_0, \mathbf{M}_0\}$ and $\{\mathbf{J}_1, \mathbf{M}_1\}$ are mathematically introduced on $\partial\Omega_0$ and $\partial\Omega_1$, respectively.

values, while the fields in Ω_1 are kept fixed, which in turn implies that the tangential fields over surface $\partial\Omega_1$ of domain Ω_1 should be unchanged, according to the uniqueness theorem [48–50]. In the equivalent problem for the scatterer, making the fields in Ω_0 trivial, $\{\mathbf{E}_0 = 0, \mathbf{H}_0 = 0\}$, leads to an arbitrariness in the ϵ_0 and μ_0 values, and a convenient choice is to take the same values as in the scatterer with $\epsilon_0 = \epsilon_1$ and $\mu_0 = \mu_1$, i.e., the full-space medium is homogeneous and isotropic. However, to maintain the tangential fields over $\partial\Omega_1$, equivalent electric and magnetic surface current densities must be introduced mathematically:

$$\mathbf{J}_1 = \hat{n}_1 \times \mathbf{H}_1, \tag{1}$$

$$\mathbf{M}_1 = -\hat{n}_1 \times \mathbf{E}_1. \tag{2}$$

Applying the Stratton-Chu equation [44] in the equivalent problem for the scatterer, the fields of interest at a point in the scatterer ($\mathbf{r} \in \Omega_1$) can be evaluated by

$$\mathbf{E}_1(\mathbf{r}) = i\omega\mu_1 \iint_{\partial\Omega_1} [\mathbf{J}_1(\mathbf{r}')G_1(\mathbf{r}, \mathbf{r}')]dS'$$

$$- \frac{1}{i\omega\epsilon_1} \iint_{\partial\Omega_1} [\nabla' \cdot \mathbf{J}_1(\mathbf{r}')\nabla G_1(\mathbf{r}, \mathbf{r}')]dS'$$

$$+ \iint_{\partial\Omega_1} [\mathbf{M}_1(\mathbf{r}') \times \nabla G_1(\mathbf{r}, \mathbf{r}')]dS', \tag{3}$$

$$\mathbf{H}_1(\mathbf{r}) = -\iint_{\partial\Omega_1} [\mathbf{J}_1(\mathbf{r}') \times \nabla G_1(\mathbf{r}, \mathbf{r}')]dS'$$

$$+ i\omega\epsilon_1 \iint_{\partial\Omega_1} [\mathbf{M}_1(\mathbf{r}')G_1(\mathbf{r}, \mathbf{r}')]dS'$$

$$- \frac{1}{i\omega\mu_1} \iint_{\partial\Omega_1} [\nabla' \cdot \mathbf{M}_1(\mathbf{r}')\nabla G_1(\mathbf{r}, \mathbf{r}')]dS', \tag{4}$$

where $G_1(\mathbf{r}, \mathbf{r}') = \exp(ik_1|\mathbf{r} - \mathbf{r}'|)/(4\pi|\mathbf{r} - \mathbf{r}'|)$ is Green's function in the whole space of the equivalent problem, and k_1 is the wave number with $k_1^2 = \epsilon_1\mu_1\omega^2$.

In the equivalent problem for free space, the fields at a point $\mathbf{r} \in \Omega_0$ can be evaluated similarly by

$$\mathbf{E}_0(\mathbf{r}) = \mathbf{E}_0^{(i)}(\mathbf{r}) + i\omega\mu_0 \iint_{\partial\Omega_0} [\mathbf{J}_0(\mathbf{r}')G_0(\mathbf{r}, \mathbf{r}')]dS'$$

$$- \frac{1}{i\omega\epsilon_0} \iint_{\partial\Omega_0} [\nabla' \cdot \mathbf{J}_0(\mathbf{r}')\nabla G_0(\mathbf{r}, \mathbf{r}')]dS'$$

$$+ \iint_{\partial\Omega_0} [\mathbf{M}_0(\mathbf{r}') \times \nabla G_0(\mathbf{r}, \mathbf{r}')]dS', \tag{5}$$

$$\mathbf{H}_0(\mathbf{r}) = \mathbf{H}_0^{(i)}(\mathbf{r}) - \iint_{\partial\Omega_0} [\mathbf{J}_0(\mathbf{r}') \times \nabla G_0(\mathbf{r}, \mathbf{r}')]dS'$$

$$+ i\omega\epsilon_0 \iint_{\partial\Omega_0} [\mathbf{M}_0(\mathbf{r}')G_0(\mathbf{r}, \mathbf{r}')]dS'$$

$$- \frac{1}{i\omega\mu_0} \iint_{\partial\Omega_0} [\nabla' \cdot \mathbf{M}_0(\mathbf{r}')\nabla G_0(\mathbf{r}, \mathbf{r}')]dS', \tag{6}$$

where the extra terms $\mathbf{E}_0^{(i)}$ and $\mathbf{H}_0^{(i)}$ account for the incident fields in Ω_0, $\mathbf{J}_0 = \hat{n}_0 \times \mathbf{H}_0$ and $\mathbf{M}_0 = -\hat{n}_0 \times \mathbf{E}_0$ are the equivalent electric and magnetic surface current densities that are introduced mathematically to maintain the tangential fields on $\partial\Omega_0$, and G_0 is the full-space Green function where the corresponding wave number is k_0 with $k_0^2 = \epsilon_0\mu_0\omega^2$.

To summarize, the fields in each region are compactly written as

$$\mathbf{E}_l = \delta_{l0}\mathbf{E}_l^{(i)} + \mathcal{D}_l\mathbf{J}_l - \mathcal{K}_l\mathbf{M}_l, \tag{7}$$

$$\mathbf{H}_l = \delta_{l0}\mathbf{H}_l^{(i)} + \mathcal{K}_l\mathbf{J}_l + \left(\frac{\epsilon_l}{\mu_l}\right)\mathcal{D}_l\mathbf{M}_l, \tag{8}$$

where δ_{l0} is the Kronecker delta and the integro-differential operators are defined as

$$\{\mathcal{D}_l\mathbf{f}(\mathbf{r}')\}(\mathbf{r}) = i\omega\mu_l \iint_{\partial\Omega_l} [\mathbf{f}(\mathbf{r}')G_l(\mathbf{r}, \mathbf{r}')]dS'$$

$$- \frac{1}{i\omega\epsilon_l} \iint_{\partial\Omega_l} [\nabla' \cdot \mathbf{f}(\mathbf{r}')\nabla G_l(\mathbf{r}, \mathbf{r}')]dS', \tag{9}$$

$$\{\mathcal{K}_l \mathbf{f}(\mathbf{r}')\}(\mathbf{r}) = - \iint_{\partial \Omega_l} [\mathbf{f}(\mathbf{r}') \times \nabla G_l(\mathbf{r}, \mathbf{r}')] dS', \quad (10)$$

with $l \in \{0, 1\}$ the domain index. Therefore, for a given point on the scatterer's surface, we have, after taking the operation $\hat{n}_l \times$ to Eqs. (7) and (8),

$$\mathbf{J}_l = \delta_{l0} \mathbf{J}_l^{(i)} + \hat{n}_l \times \left[\mathcal{K}_l \mathbf{J}_l + \left(\frac{\epsilon_l}{\mu_l} \right) \mathcal{D}_l \mathbf{M}_l \right], \quad (11)$$

$$\mathbf{M}_l = \delta_{l0} \mathbf{M}_l^{(i)} - \hat{n}_l \times [\mathcal{D}_l \mathbf{J}_l - \mathcal{K}_l \mathbf{M}_l], \quad (12)$$

where the current densities $\mathbf{M}_l^{(i)} = -\hat{n}_l \times \mathbf{E}_l^{(i)}$ and $\mathbf{J}_l^{(i)} = \hat{n}_l \times \mathbf{H}_l^{(i)}$ are due to the incident fields in free space.

The N-Müller formulation [14,43] yields stable solutions even at a low-frequency limit and it can be obtained by combining the above two equations with weights μ_l and ϵ_l, as well as taking into account the tangential field continuity at the interface (i.e., $\mathbf{J} = \mathbf{J}_0 = -\mathbf{J}_1$ and $\mathbf{M} = \mathbf{M}_0 = -\mathbf{M}_1$),

$$(\mu_0 + \mu_1)\mathbf{J} = (\mu_0 \hat{n}_0 \times \mathcal{K}_0 + \mu_1 \hat{n}_1 \times \mathcal{K}_1)\mathbf{J} + \mu_0 \mathbf{J}_0^{(i)}$$
$$+ (\epsilon_0 \hat{n}_0 \times \mathcal{D}_0 + \epsilon_1 \hat{n}_1 \times \mathcal{D}_1)\mathbf{M}, \quad (13)$$

$$(\epsilon_0 + \epsilon_1)\mathbf{M} = -(\epsilon_0 \hat{n}_0 \times \mathcal{D}_0 + \epsilon_1 \hat{n}_1 \times \mathcal{D}_1)\mathbf{J} + \epsilon_0 \mathbf{M}_0^{(i)}$$
$$+ (\mu_0 \hat{n}_0 \times \mathcal{K}_0 + \mu_1 \hat{n}_1 \times \mathcal{K}_1)\mathbf{M}. \quad (14)$$

This can be further written in the compact form

$$(\mathcal{Z}_0 + \mathcal{Z}_1)\mathbf{x} = \mathbf{b} \quad (15)$$

with the vector $\mathbf{x} = [\mathbf{J}, \mathbf{M}]$ containing the equivalent surface current densities to be solved with the knowledge of the tangential components of the incident fields $\mathbf{b} = [\mu_0 \mathbf{J}_0^{(i)}, \epsilon_0 \mathbf{M}_0^{(i)}]$. We also invoked the matrix notation

$$\mathcal{Z}_l = \begin{bmatrix} \mu_l & 0 \\ 0 & \epsilon_l \end{bmatrix} + \hat{n}_l \times \begin{bmatrix} -\mu_l \mathcal{K}_l & -\epsilon_l \mathcal{D}_l \\ \epsilon_l \mathcal{D}_l & -\mu_l \mathcal{K}_l \end{bmatrix}. \quad (16)$$

B. Eigenmodes

In the absence of an incident field $\mathbf{b} = 0$ holds, and the system may have nontrivial solutions at certain frequencies (usually complex valued), each of them being associated with a resonance mode of the optical system [14,22,23,33]. The resonance frequencies are generally isolated poles in the complex frequency plane. At a fixed real-valued frequency, however, we can find the scatterer's eigenmodes

[14,33] by solving

$$(\mathcal{Z}_0 + \mathcal{Z}_1)\mathbf{x}^{(m)} = \lambda^{(m)} \mathbf{x}^{(m)}. \quad (17)$$

Rewriting Eq. (17) in the form of Eq. (15), we have

$$(\mathcal{Z}_0 + \mathcal{Z}_1) \frac{\mathbf{x}^{(m)}}{\lambda^{(m)}} = \mathbf{x}^{(m)}. \quad (18)$$

Comparing Eqs. (15) and (18), it is seen that an incident field $\mathbf{x}^{(m)}$, which is the mth eigenmode, yields a scattered field $\mathbf{x}^{(m)}/\lambda^{(m)}$. In addition, an eigenmode associated with a smaller eigenvalue $\lambda^{(m)}$ yields a stronger scattered field $\mathbf{x}^{(m)}/\lambda^{(m)}$ when the scatterer is excited by an incident field of $\mathbf{x}^{(m)}$. Therefore, $1/\lambda^{(m)}$ can be viewed as the *scattering strength* associated with the mth eigenmode.

For a given eigenmode $\mathbf{x}^{(m)} = [\mathbf{J}^{(m)}, \mathbf{M}^{(m)}]$, we can calculate its electric and magnetic fields $\mathbf{E}^{(m)}$ and $\mathbf{H}^{(m)}$ everywhere, by invoking Eqs. (7) and (8),

$$\mathbf{E}_l^{(m)}(\mathbf{r}) = \mathcal{D}_l \mathbf{J}_l^{(m)} - \mathcal{K}_l \mathbf{M}_l^{(m)}, \quad (19)$$

$$\mathbf{H}_l^{(m)}(\mathbf{r}) = \mathcal{K}_l \mathbf{J}_l^{(m)} + \left(\frac{\epsilon_l}{\mu_l} \right) \mathcal{D}_l \mathbf{M}_l^{(m)}, \quad (20)$$

where the subscript l denotes the domain where the field at a given point $\mathbf{r} \in \Omega_l$ is evaluated. Far fields can be obtained by making the far-field approximation to the above equations, i.e., the scalar Green function's exponential term is factored into a stationary term and an oscillatory term [48].

C. Silicon nanodisks

1. Monomer

For the purpose of demonstrating the eigenmodes in all-dielectric nanostructures, we study a silicon nanodisk in air. The nanodisk has a thickness of $h = 220$ nm and a diameter of $d = 480$ nm and it is situated at the origin of the Cartesian coordinate system, as shown in the inset of Fig. 2. The eigenvalues are solved in the wavelength range [850, 1750] nm and the twelve dominant eigenmodes are shown in Fig. 2. The eigenmodes sharing the same $1/|\lambda^{(m)}|$ values are degenerate modes, whose fields are of identical distribution but distinctively oriented. Treating the degenerate eigenmodes together, we observe seven unique eigenmodes showing different resonance responses. Mode M1 (M2) has a resonance wavelength around 1330 nm. Mode M3 (M4) does not show clear resonance as its scattering strength value is flattened over a relatively wide spectral range. Mode M5 is a nondegenerate mode showing resonance scattering at 1675 nm. Mode M6 (M7) has a relatively weak resonance near 1200 nm. Resonances at two shorter wavelengths 970 and 1070 nm are associated with modes M8 (M9) and M11 (M12), respectively.
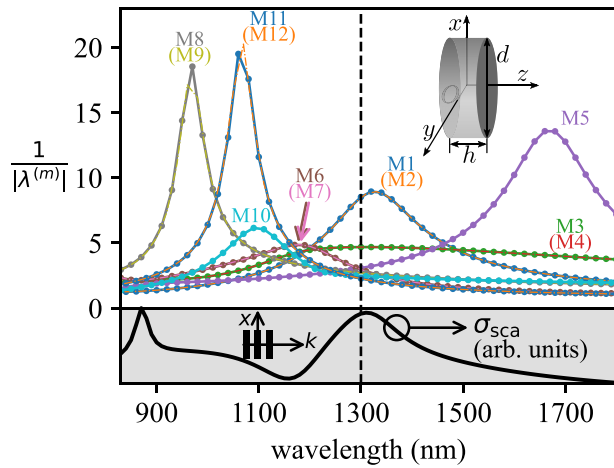
FIG. 2. The reciprocal of absolute eigenvalues $1/|\lambda^{(m)}|$ of the first twelve leading eigenmodes in a silicon nanodisk in air as a function of the wavelength. The nanodisk has a thickness of $h = 220$ nm and a diameter of $d = 480$ nm and it is situated at the origin of the Cartesian coordinate system (shown in the inset). Each eigenmode is labeled M1 to M12, respectively. In the case of degeneracy, one of the two degenerate eigenmode labels is bracketed. These twelve eigenmodes are ranked by their $1/|\lambda^{(m)}|$ values at wavelength 1300 nm (the dashed line). In the filled gray area below the zero horizontal line, the scattering cross section (σ_{sca}) in arbitrary units is shown as the black curve when the nanodisk is illuminated by an x-polarized plane wave.

Finally, mode M10 is weakly resonant across the entire spectral range of interest. At this point, we emphasize that the resonance wavelengths determined from the scattering strength values are the genuine resonance wavelengths that do not depend on any external incident field.

The calculated scattering cross section in arbitrary units from the nanodisk when illuminated by an x-polarized plane wave propagating along the z axis is shown as the black curve in the filled gray area. The scattering cross section under the illumination of a linearly polarized plane wave shows a resonance peak near 1300 nm (the dashed line), at which we have ranked the aforementioned eigenmodes by their scattering strength values $1/|\lambda^{(m)}|$. It is worth emphasizing that the excitation field is a plane wave, which effectively matches and thus excites mode M1 (M2) but no other eigenmodes. Therefore, the excitations to other eigenmodes by the plane-wave field are weak and the resonances associated with other eigenmodes do not manifest themselves in the scattering cross-section spectrum. This also implies that we need to design the incident vector fields to excite eigenmodes that are not observable under a plane-wave excitation.

The corresponding near and far fields of the eigenmodes at wavelength 1300 nm are inspected in Fig. 3, and only one of degenerate eigenmodes, if any, is considered. For visualizing both the electric and magnetic near fields
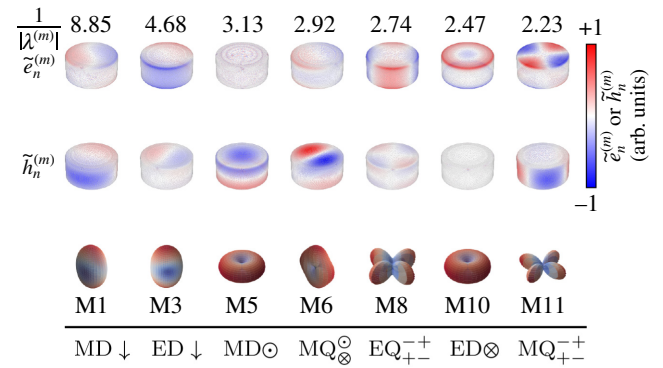

FIG. 3. Near fields including the instantaneous electric $\widetilde{e}_n^{(m)}$ (first row) and magnetic $\widetilde{h}_n^{(m)}$ (second row) fields on the nanodisk's surface, and far fields (third row) of a few leading eigenmodes at wavelength 1300 nm for the same silicon nanodisk as in Fig. 2. Each column is for one eigenmode with its $1/|\lambda^{(m)}|$ value given at the top.

with balanced amplitudes [51,52], we introduce the scaled instantaneous electric and magnetic fields as

$$\widetilde{\mathbf{e}}^{(m)}(\mathbf{r}, t) = \sqrt{\epsilon_0}\mathrm{Re}\{\mathbf{E}_0^{(m)}(\mathbf{r})e^{-i\omega t}\} \qquad (21)$$

and

$$\widetilde{\mathbf{h}}^{(m)}(\mathbf{r}, t) = \sqrt{\mu_0}\mathrm{Re}\{\mathbf{H}_0^{(m)}(\mathbf{r})e^{-i\omega t}\}, \qquad (22)$$

respectively, where t is a time instant, $r \in \partial\Omega_0$, and $\mathrm{Re}\{\cdot\}$ denotes the real part. Note that the instantaneous electric and magnetic energy densities will be written as half of the square of the scaled instantaneous electric and magnetic fields, i.e., $w_e = [\widetilde{\mathbf{e}}^{(m)}]^2/2$ and $w_h = [\widetilde{\mathbf{h}}^{(m)}]^2/2$, respectively [53]. On the nanodisk's surface, the normal component of the scaled field $\widetilde{e}_n^{(m)} = \hat{n} \cdot \widetilde{\mathbf{e}}^{(m)}$ at time instant t_0 is shown for the eigenmode's electric field, and the magnetic counterpart is shown by $\widetilde{h}_n^{(m)} = \hat{n} \cdot \widetilde{\mathbf{h}}^{(m)}$ at a $\pi/2$-phase-delayed time instant t_0' with $\omega(t_0' - t_0) = \pi/2$, at which time the magnetic field generally reaches its maximum.

For a nanodisk monomer, the surface normal near fields $\widetilde{e}_n^{(m)}$ and $\widetilde{h}_n^{(m)}$ are shown in the first and second rows, respectively, of Fig. 3, and the eigenmodes' corresponding far fields are shown in the third row. Each column displays the near and far fields for one eigenmode with its $1/|\lambda^{(m)}|$ value given at the top of each column. As can be seen from the radiation patterns, eigenmode M1 is a dipolar mode showing a typical doughnut shape. Furthermore, the near fields show that mode M1 is of magnetic dipole with a strong transverse magnetic field, which is denoted MD $\downarrow$ with the down arrow representing a field vector in the transverse plane. In the second column, M3 has a smaller $1/|\lambda^{(3)}|$ value, and the radiation pattern and near fields indicate that it is an electric dipolar mode, which is denoted ED $\downarrow$. Under a similar observation, it turns out that M5 is
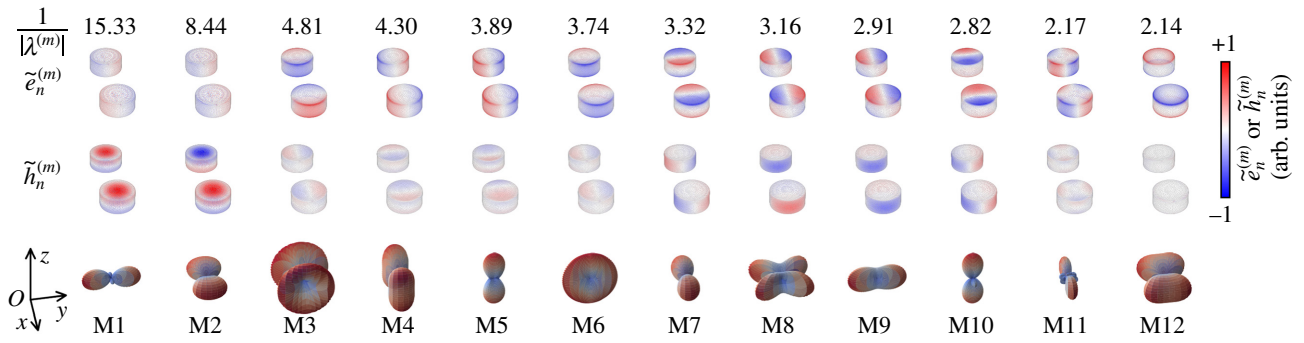
FIG. 4.   Near and far fields of eigenmodes M1–12 at 1600 nm in a dimer consisting of two identical silicon nanodisks (same as in Fig. 2) with a center-to-center gap of 480 nm.

a magnetic dipolar mode that differs from M1 by a dominating z magnetic field component, denoted MD⊙ with "⊙" implying a field vector along the positive z axis. Mode M6 emerges from a magnetic quadrupole lying in a plane parallel to the z axis, which is denoted MQ$_\otimes^\odot$, where "⊗" implies a field vector along the negative z axis that is π out of phase with respect to the "⊙" field vector. Mode M6 is distinct from another magnetic quadrupolar mode M11 where the quadrupole resides in the x-y plane, denoted MQ$_{+-}^{-+}$, where "−" and "+" represent negative and positive field extremes, respectively. An electric quadrupolar mode, denoted EQ$_{+-}^{-+}$, is observed in mode M8. Finally, mode M10 is an electric dipolar mode, denoted ED⊗, with a dominating electric z-field component.

2. Oligomers

In silicon nanodisk ensembles, i.e., nanodisk oligomers consisting of several nanodisks, further local field enhancements and mode hybridization happen via optical near-field coupling. Here, we also investigate the near fields of a few leading eigenmodes in nanostructure assemblies that consist of two and four identical nanodisks, the so-called nanodisk dimer and tetramer, respectively.

The eigenmodes' near and far fields of a dimer at 1600 nm are shown in Fig. 4. Modes M1 and M2 are hybridized from two longitudinal magnetic dipolar modes, where the former arises from two parallel longitudinal MD modes, and the latter is due to two antiparallel longitudinal MD modes. Transversal magnetic modes include M7–10 that are due to the subsequent hybridization of two y-oriented antiparallel transversal MD modes, two x-oriented antiparallel transversal MD modes, two x-oriented parallel transversal MD modes, and two y-oriented parallel transversal MD modes. The electric counterparts of magnetic modes M7–10 are observed in transversal electric modes M3–6. In detail, M3 is hybridized from two x-oriented antiparallel transversal ED modes, M4 is due to two y-oriented antiparallel transversal ED modes, M5 arises from two y-oriented parallel transversal ED modes, and two x-oriented parallel transversal ED modes yield hybrid mode M6. Additionally, two transversal EQ modes, when arranged differently, hybridize into two distinct modes M11 and M12.

In a tetramer, modes hybridization leads to more types of eigenmodes as a result of more near-field coupling channels. Near and far fields of twelve eigenmodes with leading scattering strengths $1/|\lambda^{(m)}|$ at 1600 nm are shown in Fig. 5. It is clear that mode M1 arises from four
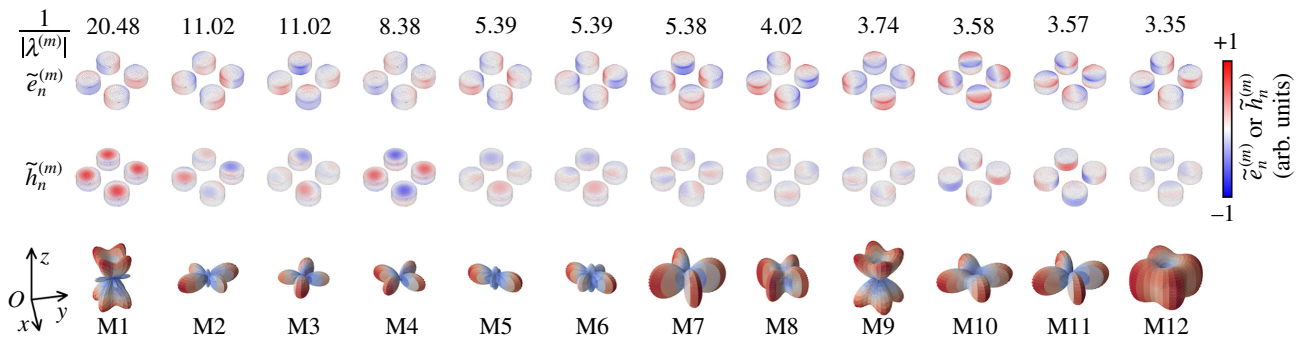

FIG. 5.   Near and far fields of a few leading eigenmodes M1–12 at 1600 nm in a tetramer consisting of four identical silicon nanodisks (same as in Fig. 2) with the center-to-center gap of two opposite neighboring nanodisks being 480 nm.

parallel longitudinal magnetic modes, whereas the longitudinal magnetic fields in every two neighboring nanodisks in M4 are π out of phase with respect to each other. Modes hybridized from four transversal ED modes contain M7 where the ED moments are π out of phase along the radial direction between every two neighboring elements, M8 where the ED moments are π out of phase between every two neighboring elements along the azimuthal direction, M9 with radially in-phase ED moment in each element, and M12 with in-phase ED moment in each nanodisk along the azimuthal direction. Hybridization of transversal MDs leads to mode M10 with the MD moments being along the radial direction and π out of phase between the neighboring nanodisks and M11 with azimuthally in-phase MD moments. In addition, M10 and M11 are the magnetic counterparts of modes M7 and M12. It is not straightforward to gain physical insight into the origins of hybridized modes M2, M3, M5, and M6 without future investigation, and determining the near-field coupling channels for these modes is less important and not further discussed in this work.

As can be seen from Figs. 3–5, an eigenmode is associated with its unique near-field distribution and far-field radiation pattern. Tuning light scattered from nano-objects via the selective excitation(s) of a certain eigenmode(s) could serve as a basic building block for engineering the optical near and/or far field(s). It is therefore vital to design an excitation field that highly overlaps with the field of a specific eigenmode for an effective mode excitation, i.e., mode-matching field.

III. INVERSE DESIGN VIA TIGHT FOCUSING

As shown in Sec. II, the eigenmode's near fields in optical nanoantennas are genuinely localized and vectorial at the nanoscale. A potential free-space approach to design such a 3D mode-matching vector field is to shape a focal field (that matches with the eigenmode's near field in optical nanoantennas that are to be placed in the vicinity of the focal point), by focusing a paraxial vector beam in an aplanatic system with high NA. In previous work, tightly focused cylindrical vector beams of fundamental order, i.e., radially and azimuthally polarized beams [17,18], and of higher-order polarization states [21] have been used for optical excitation of dark and hybrid modes in optical nanoantennas. In these forward design approaches, the focused field has a predefined field profile and only a qualitative match of polarization distribution between the hybrid modes and the excitation fields was considered. To design the best possible focal field that quantitatively matches with the eigenmode's field, it is necessary to solve an inverse problem where the needed paraxial vector beam to form a desired focal field after tight focusing is to be sought.

A. Tight focusing: forward and backward propagation

Before proceeding, we first review the forward and backward propagation in tight focusing of vector beams. In detail, we choose a coordinate system such that the middle plane of the nanostructure(s) coincides with the focal plane Γ (where $z = 0$), as shown in Fig. 6. For a given eigenmode, its near field $\{\mathbf{E}^{(m)}, \mathbf{H}^{(m)}\}$ is evaluated from Eqs. (19) and (20) in the focal plane Γ. For the eigenmode's electric near field to be mode matched, a 3D vector field $\mathbf{E}_f$ is designed via the tight focusing of a paraxial vector beam $\mathbf{E}_{\mathrm{inc}}$ at the pupil plane (which is mapped to $\mathbf{E}_\infty$ on the reference sphere).

The tight focusing process is governed by the Richards-Wolf formalism [54–56], and the focal field at a point
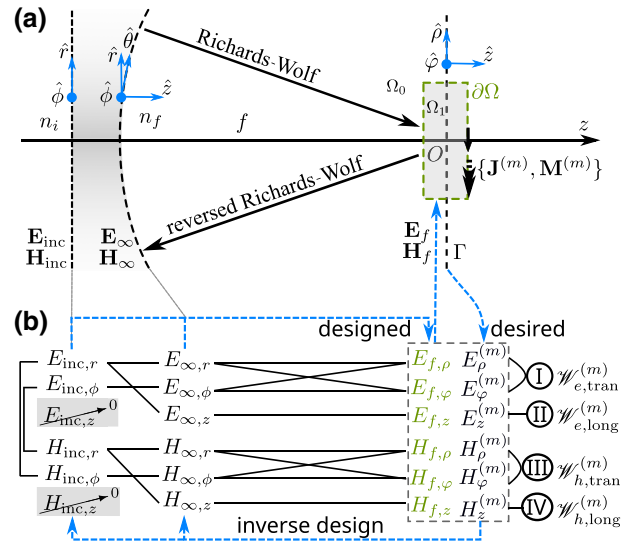

FIG. 6. The inverse design scheme of focused vector beams with effective mode-matching profiles. (a) The forward and backward tight focusing processes. The nanostructure's surface is denoted $\partial\Omega$, where the mth eigenmode's equivalent surface current density $\{\mathbf{J}^{(m)}, \mathbf{M}^{(m)}\}$ resides. The middle symmetry plane Γ coincides with the focal plane ($z = 0$ plane), where the modal field $\{\mathbf{E}^{(m)}, \mathbf{H}^{(m)}\}$ is evaluated. The reference field $\mathbf{E}_\infty$ resides on the reference spherical cap with a radius equal to the focal length f. The incident beam is denoted $\mathbf{E}_{\mathrm{inc}}$. The nanostructure occupies domain Ω_1, whereas the free space is denoted Ω_0 with a refractive index of n_f. The index before the reference sphere is denoted n_i. (b) The inverse design flowchart. Solid black lines indicate connections between the associated field components. The modal field $\{\mathbf{E}^{(m)}, \mathbf{H}^{(m)}\}$ on the focal plane Γ serves as the desired focal field, from which the pupil field $\{\mathbf{E}_{\mathrm{inc}}, \mathbf{H}_{\mathrm{inc}}\}$ and thus the focal field $\{\mathbf{E}_f, \mathbf{H}_f\}$ are designed. Desired focal fields are categorized into four groups by the time-averaged energies stored in the transverse and longitudinal, electric and magnetic fields in the focal plane, denoted $\mathscr{W}_{e,\mathrm{tran}}^{(m)}$, $\mathscr{W}_{e,\mathrm{long}}^{(m)}$, $\mathscr{W}_{h,\mathrm{tran}}^{(m)}$, and $\mathscr{W}_{h,\mathrm{long}}^{(m)}$, respectively, in Eqs. (30)–(33).

$\mathbf{R}_f = (\rho, \varphi, z)$ in cylindrical coordinates reads

$$\mathbf{E}_f(\rho, \varphi, z) = \frac{-ik}{2\pi} \int_0^{\theta_{\max}} \int_0^{2\pi} \mathbf{E}_\infty(\theta, \phi) e^{i\mathbf{k}\cdot\mathbf{R}_f} \sin\theta \, d\phi \, d\theta,$$ (23)

where the wave vector $\mathbf{k}$ is along the direction from a point on the reference sphere (ϕ, θ) to the focal point and the corresponding wave number $k = k_0$ (i.e., the same value as in domain Ω_0). The maximum converging angle $\theta_{\max} = \arcsin(\mathrm{NA}/n_f)$ is limited by the NA of the aplanatic system, where n_f is the refractive index after the reference sphere. We can rewrite Eq. (23) in Fourier transform in Cartesian coordinates as [57,58]

$$\mathbf{E}_f(x, y, z) = \frac{-ik}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \left[\frac{\mathbf{E}_\infty(k_x, k_y) e^{ikz\cos\theta}}{k^2 \cos\theta} \right]$$
$$\times e^{-ik_x x} e^{-ik_y y} dk_x dk_y, \quad (24)$$

where $k_x = -k\sin\theta\cos\phi$ and $k_y = -k\sin\theta\sin\phi$ are the x and y components of the wave number, respectively. Here, the surface integral runs from $-\infty$ to $+\infty$, as the field $\mathbf{E}_\infty$ on the reference sphere is zero outside of the NA. In order to obtain the necessary paraxial vector beam or the field $\mathbf{E}_\infty$ on the reference sphere to generate the desired 3D mode-matching vector field $\mathbf{E}_f$, we perform the inverse Fourier transform of Eq. (24) except $\mathbf{E}_f$ is replaced by the eigenmode's near field in the focal plane $\mathbf{E}^{(m)}|_\Gamma$,

$$\frac{\mathbf{E}_\infty(k_x, k_y) e^{ikz\cos\theta}}{\cos\theta} = \frac{ik}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \mathbf{E}^{(m)}|_\Gamma \, e^{ik_x x} e^{ik_y y} dx dy,$$ (25)

which can be considered as the governing equation for a reversed Richards-Wolf tight focusing process.

Ideally, the paraxial input vector beam is inversely restored from the reference field [21,59]

$$\mathbf{E}_{\mathrm{inc}}(r, \phi) = \mathbf{L}^{-1}(\theta) \mathbf{E}_\infty(\theta, \phi),$$ (26)

where the radius $r = f\sin\theta$ with f being the focal length and the mapping matrix in cylindrical coordinates is [21]

$$\mathbf{L}^{-1}(\theta) = \frac{1}{f\sqrt{\cos\theta}} \sqrt{\frac{n_f}{n_i}} \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix}, \quad (27)$$

where n_i is the refractive index before the reference sphere. Above, we followed Richards and Wolf [55] for the refraction at the aplanatic lens; the reference field amplitude vector reads $\mathbf{E}_\infty(\theta, \phi) = f\sqrt{n_i/n_f}\sqrt{\cos\theta}[E_{\mathrm{inc},r}\hat{\theta} + E_{\mathrm{inc},\phi}\hat{\phi} + E_{\mathrm{inc},z}(\hat{\theta} \times \hat{\phi})]$.

B. Degrees of freedom

The associated magnetic fields are governed by similar formulas as in Eqs. (23)–(26) for the electric counterparts, but, for brevity, we do not repeat them here. However, note that in the Richards-Wolf formalism the magnetic fields are linked to electric fields by revoking the relations $\mathbf{H}_{\mathrm{inc}} = 1/Z_i(\mathbf{k}_i/k_i) \times \mathbf{E}_{\mathrm{inc}}$ and thus $\mathbf{H}_\infty = 1/Z_f(\mathbf{k}_f/k_f) \times \mathbf{E}_\infty$ [54] as a result of a beamlike characteristic of the vector fields on the pupil plane and on the reference sphere. Here, Z_i and Z_f are the associated wave impedances. This implies that only two components in the pupil vector field are completely independent. In cylindrical coordinates, four possible combinations of two field components are $(E_{\mathrm{inc},r}, E_{\mathrm{inc},\phi})$, $(H_{\mathrm{inc},r}, H_{\mathrm{inc},\phi})$, $(E_{\mathrm{inc},r}, H_{\mathrm{inc},r})$, and $(E_{\mathrm{inc},\phi}, H_{\mathrm{inc},\phi})$, and choosing any pair will fully determine the vector field on the pupil plane, then on the reference sphere, and finally in the focal region.

The fact that the incident vector beam is paraxial ($E_{\mathrm{inc},z} \approx 0$ and $H_{\mathrm{inc},z} \approx 0$) imposes a restriction on the radial and longitudinal components of the reference electric [58] and magnetic fields via the relations

$$-\sin\theta E_{\infty,r}(\theta, \phi) + \cos\theta E_{\infty,z}(\theta, \phi) = 0,$$ (28)

$$-\sin\theta H_{\infty,r}(\theta, \phi) + \cos\theta H_{\infty,z}(\theta, \phi) = 0.$$ (29)

As a consequence, the focused electric field $\mathbf{E}_f$ bears all three components that are nevertheless not completely independent [39]. In other words, all three components of the electric focal field are determined uniquely by two components of the beamlike pupil field $\mathbf{E}_{\mathrm{inc}}$, as seen in Fig. 6(b) where the black solid lines indicate the dependencies of field components on different planes. In order to inversely design a focal field that is truly generated by a beamlike vector pupil field, inherent dependencies of different field components of the desired focal field should be carefully investigated.

C. Inverse design strategies

To inversely design a focal field that potentially matches with an eigenmode's field, i.e., the desired focal field, in principle all the electric and magnetic modal field components should be taken into account. However, bearing in mind the degrees of freedom discussed in the previous section, an inconsistent pupil field could result from a backward propagation that ignores the inherent dependencies of all components [39] in the desired focal field. In contrast, one has to choose appropriate components in the desired focal field that will unambiguously determine the beamlike pupil field and thus the designed focal field. It can be seen from Figs. 3–5 that only some of the electric and magnetic field components are dominant in an eigenmode's near field. For instance, mode M10 of a single

nanodisk in Fig. 3 shows strong longitudinal electric field component. Moreover, in Fig. 4 modes M3–6 of a nanodisk dimer are clearly dominated by their transverse electric and longitudinal magnetic field components, whereas the transverse magnetic and longitudinal electric field components are dominant in modes M7–10 of a nanodisk dimer.

For a more quantitative analysis, we categorize the desired focal fields according to the time-averaged energies associated with the transverse and longitudinal electric and magnetic fields in the focal plane, i.e., the following surface integrals of the energy densities [60]:

$$\mathscr{W}_{e,\text{tran}}^{(m)} = \frac{1}{2}\epsilon_0 \int_\Gamma (|E_\rho^{(m)}|^2 + |E_\varphi^{(m)}|^2)dS, \qquad (30)$$

$$\mathscr{W}_{e,\text{long}}^{(m)} = \frac{1}{2}\epsilon_0 \int_\Gamma |E_z^{(m)}|^2 dS, \qquad (31)$$

$$\mathscr{W}_{h,\text{tran}}^{(m)} = \frac{1}{2}\mu_0 \int_\Gamma (|H_\rho^{(m)}|^2 + |H_\varphi^{(m)}|^2)dS, \qquad (32)$$

$$\mathscr{W}_{h,\text{long}}^{(m)} = \frac{1}{2}\mu_0 \int_\Gamma |H_z^{(m)}|^2 dS. \qquad (33)$$

In principle, the surface integration is performed over an infinite focal plane Γ, but in numerical evaluation we have to truncate the integration surface in an area where the desired focal field remains significant and discard the field elsewhere.

In group ①, the focal-plane energy associated with the transverse electric field is larger than that of the longitudinal electric field component as well as both focal-plane energies related to the transverse and longitudinal magnetic field components. In group ②, the focal-plane energy of the longitudinal electric field overwhelms the other three types of energies. The magnetic counterparts of groups ① and ② are groups ③ and ④, respectively. Concisely, we may write

$$\max(\mathscr{W}_{e,\text{tran}}^{(m)}, \mathscr{W}_{e,\text{long}}^{(m)}, \mathscr{W}_{h,\text{tran}}^{(m)}, \mathscr{W}_{h,\text{long}}^{(m)})$$

$$= \begin{cases} \mathscr{W}_{e,\text{tran}}^{(m)}, & \text{in group ①}, \\ \mathscr{W}_{e,\text{long}}^{(m)}, & \text{in group ②}, \\ \mathscr{W}_{h,\text{tran}}^{(m)}, & \text{in group ③}, \\ \mathscr{W}_{h,\text{long}}^{(m)}, & \text{in group ④}. \end{cases} \qquad (34)$$

For eigenmodes belonging to group ①, the corresponding transverse electric field is chosen as the candidate component of the desired focal field, upon which the radial and azimuthal components of the reference field $(E_{\infty,r}, E_{\infty,\phi})$ and thus the pupil field $(E_{\text{inc},r}, E_{\text{inc},\phi})$ are determined. The longitudinal electric focal field is, however, a free parameter so that the longitudinal electric field component on the reference sphere $E_{\infty,z}$ is determined from the radial component $E_{\infty,r}$ according to the restriction in Eq. (28). In this

regard, the pupil vector field is guaranteed to be beamlike, and the designed focal field including the magnetic part is fully determined by the pair of field components $(E_{\text{inc},r}, E_{\text{inc},\phi})$.

For eigenmodes in group ②, the longitudinal electric field component of the desired focal field $E_z^{(m)}$ determines the longitudinal electric reference field $E_{\infty,z}$, then the radial electric pupil field $E_{\text{inc},r}$ using the restriction in Eq. (28) where the radial electric reference field is a free parameter. What remains to be determined is either the azimuthal electric or the radial magnetic pupil field. But, the azimuthal electric pupil field is only determined by the azimuthal electric reference field, which in turn depends on the radial and azimuthal desired electric focal fields, both of which are usually trivial for an eigenmode showing strong longitudinal electric field. On the other hand, a strong longitudinal electric field component of the desired focal field is usually accompanied by a slightly less strong transverse magnetic field component $H_\rho^{(m)}$ and/or $H_\varphi^{(m)}$ from which the radial magnetic reference field $H_{\infty,r}$ and then the radial magnetic pupil field $H_{\text{inc},r}$ can be determined. The magnetic counterparts of the aforementioned electric design strategies apply to groups ③ and ④. At this point, it is worth pointing out that the restriction on the field components $E_{\infty,r}$ and $E_{\infty,z}$, as well as $H_{\infty,r}$ and $H_{\infty,z}$, poses a challenge to design a 3D focused vector beam, whose electric and magnetic field components perfectly match the eigenmode's near field.

Besides the constraint imposed by the beamlike characteristics in the pupil field, another limitation in the inverse design of a focal field is due to the fact that the reference field is band limited, i.e., spatial frequencies that are larger than the free-space wave number are evanescent and cannot be backward propagated onto the reference sphere. This is associated with the fact that the tight focusing process relies on free-space beam propagation, and the focal field is actually the far field propagated from the secondary field on the reference sphere. In addition, backpropagated waves are further limited by the NA of the aplanatic system.

An example eigenmode belonging to group ① is mode M2 in the nanodisk dimer, and its scaled electric and magnetic fields at different stages of the inverse design are shown in Fig. 7. The first and second columns show all six field components of the desired focal field in real space and spatial frequency **k** space, respectively. The inversely designed pupil field's components are shown in the third column, and the designed focal field's components can be seen in the fourth column. Both the radial and azimuthal scaled electric field components of the designed focal field effectively match with the desired eigenmode's near field, even though the desired focal field is so locally confined that a considerable amount of light is at high spatial frequencies outside the k_0 circle and has to be discarded in the free-space backward propagation of the inverse design.
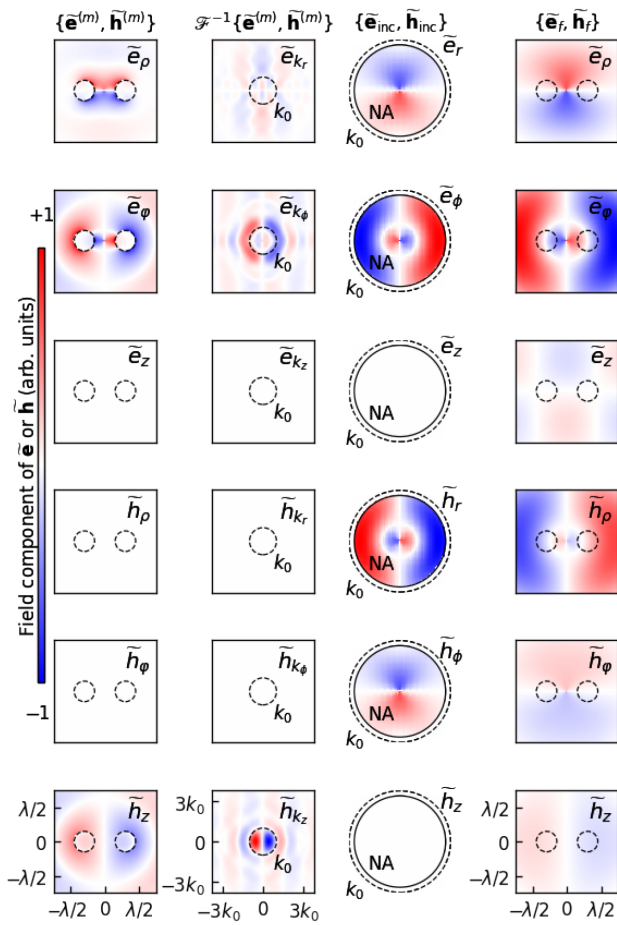
FIG. 7. Scaled electric and magnetic field components [defined in Eqs. (21) and (22)] at different stages in the inverse design for M2 in the nanodisk dimer whose eigenmodes are shown in Fig. 4. The dashed circles represent the nanodisk's outlines in the desired (first column) and designed (fourth column) focal fields, or the threshold of maximum free-space wave number k_0 in **k** space of the desired focal field (second column) and the designed pupil field (third column). The solid circle in the designed pupil field indicates the wave number limited by NA with a value of 0.9 used in this inverse design.

Imposed by the beamlike characteristics of the designed pupil field, the azimuthal scaled magnetic field is identical to the radial scaled electric field, whereas the radial scaled magnetic field is π-phase shifted with respect to the azimuthal scaled electric field. Four other field components are also present as a result of tight focusing of the beamlike pupil field and they are dependent on the radial and azimuthal electric field components [see Fig. 6(b) for the connections of field components]. In particular, the designed longitudinal magnetic field is π out of phase with respect to the desired longitudinal magnetic field. Nevertheless, as will be shown with the expansion coefficients in the next section, this designed focal field still effectively excites mode M2 in the nanodisk dimer.
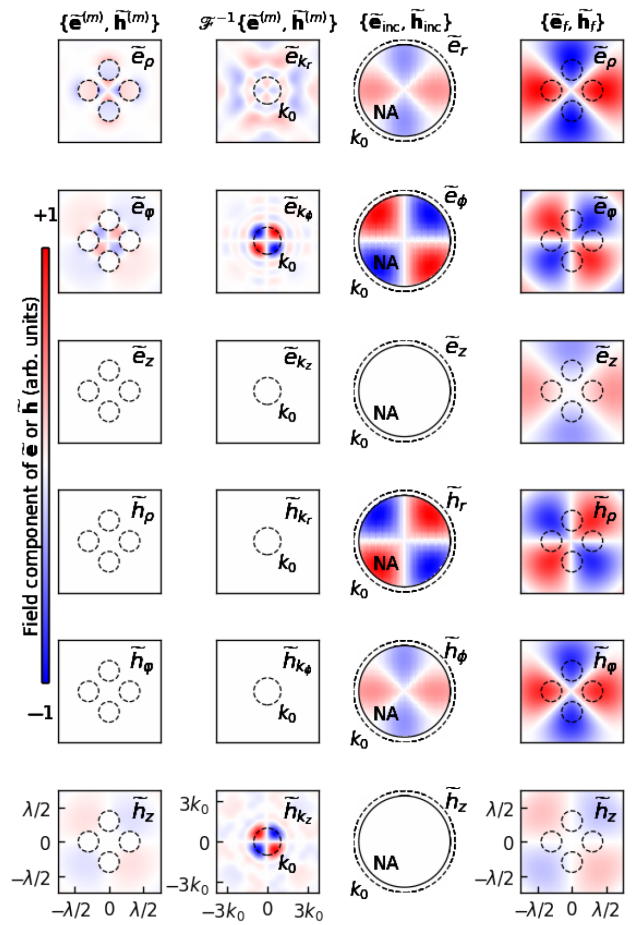
FIG. 8. Scaled electric and magnetic field components in the inverse design for M7 in the nanodisk tetramer (similar to Fig. 7) whose eigenmodes are shown in Fig. 5.

Another example is shown in Fig. 8 for inversely designing a focal field to match with eigenmode M7 in the nanodisk tetramer. An effective match between the designed and desired focal fields is achieved in radial and azimuthal electric field components, as well as the longitudinal magnetic field component. Again, as will be shown with the expansion coefficients in the next section, this designed focal field can effectively and exclusively excite mode M7 in the nanodisk tetramer. In this case, it seems that dropping waves with spatial frequencies higher than those limited by NA has a trivial effect on the excitation efficiency. A full list of scaled field components in the inverse design of the focal field for all twelve eigenmodes in the nanodisk monomer, dimer, and tetramer is compiled in the Supplemental Material [61].

From a practical point of view, it is worth mentioning that we are designing incident vector fields rather than nanostructures. Provided that fabricated structures suffer from distortions from the ideal ones due to fabrication imperfections, SEM imaging techniques would measure and reveal most types of deviations, which, in turn, can

be taken into account in our inverse design of the incident vector fields from the very beginning. In other words, we may perform an eigenmode analysis and an inverse design to the realistic structure directly obtained from SEM imaging rather than to the ideal one, and the impact of fabrication errors is always diminished.

IV. MODE EXPANSION COEFFICIENTS

Following the analysis of eigenmodes in nano-objects (Sec. II) and the inverse design of a tightly focused beam (Sec. III), it is important to see how well a designed focal field matches with an eigenmode's near field. A figure of merit is the modal expansion coefficient that characterizes the overlap between the inversely designed vector beam and the eigenmode's near field. Under the SIE formulation, the modal expansion coefficient can be evaluated via the inner product of two light fields over the nanostructures' surfaces $\partial\Omega_0$ [14],

$$a^{(m,n)} = \frac{\iint_{\partial\Omega_0} \mathbf{x}_L^{(n)} \cdot \mathbf{b}^{(m)} dS}{\iint_{\partial\Omega_0} \mathbf{x}_L^{(n)} \cdot \mathbf{x}^{(n)} dS}, \qquad (35)$$

where $\mathbf{x}_L^{(n)}$ and $\mathbf{x}^{(n)}$ are the left and right eigenvectors associated with the nth eigenmode with an eigenvalue $\lambda^{(n)}$, and $\mathbf{b}^{(m)}$ is the vector representation of the focal field inversely designed from the mth eigenmode and evaluated on the surface of the nanodisk(s) $\partial\Omega_0$. In detail, the left eigenvector is defined by, in comparison with Eq. (17),

$$\mathbf{x}_L^{(n)}(\mathcal{Z}_0 + \mathcal{Z}_1) = \lambda^{(n)} \mathbf{x}_L^{(n)}, \qquad (36)$$

and the vector of the designed focal field is

$$\mathbf{b}^{(m)} = [\mu_0 \mathbf{J}_f^{(m)}, \epsilon_0 \mathbf{M}_f^{(m)}], \qquad (37)$$

where $\mathbf{J}_f^{(m)} = \hat{n}_0 \times \mathbf{H}_f^{(m)}$ and $\mathbf{M}_f^{(m)} = -\hat{n}_0 \times \mathbf{E}_f^{(m)}$ are the electric and magnetic surface current densities arising from the designed focal field over surface $\partial\Omega_0$.

The eigenmode's near field is solved directly in Eq. (17) over surface $\partial\Omega_0$. The designed focal field on $\partial\Omega_0$ can be evaluated via the surface double integral in Eq. (23), where the reference field $\mathbf{E}_\infty$ in the integrand is calculated using Eq. (25) by fast Fourier transform [62]. As a result, the reference field $\mathbf{E}_\infty$ is a 2D matrix on a square grid, and we need to use Eq. (23) rather than Eq. (24) for the focal field calculation on $\partial\Omega_0$. Here, the double integral is implemented through the 2D numerical integration based on the trapezoidal rule [63].

The expansion coefficients for the nanodisk monomer, dimer, and tetramer are visualized in three colormap matrices in Figs. 9(a)–9(c), respectively. The mth focal field $\mathbf{E}_f^{(m)}$ is inversely designed from the corresponding mth eigenmode's near field in the focal plane. For the monomer, the designed focal fields $\mathbf{E}_f^{(3)}$, $\mathbf{E}_f^{(4)}$, $\mathbf{E}_f^{(5)}$, $\mathbf{E}_f^{(8)}$, $\mathbf{E}_f^{(9)}$, and $\mathbf{E}_f^{(10)}$ largely overlap with the eigenmodes from which they are designed, respectively. In other words, these eigenmodes can be effectively mode matched by tightly focused vector beams via our inverse design approach. However, the designed focal fields $\mathbf{E}_f^{(m)}$ with $m = 1, 2, 6, 7, 11$, and 12 do not effectively match the corresponding desired eigenmodes and instead other eigenmodes are strongly excited. In the nanodisk dimer case, the effective mode-matching focal fields can be obtained for the eigenmodes $\mathbf{E}^{(n)}$ with $n = 1$ to 3 and 10, but the focal fields designed from other eigenmodes jump to near fields that do not effectively match with the corresponding eigenmodes. For the nanodisk tetramer, the effective mode-matching focal fields are obtained for eigenmodes $\mathbf{E}^{(n)}$ with $n = 1, 2, 5, 6, 7, 9$, and 10. However, the designed focal fields $\mathbf{E}_f^{(3)}$, $\mathbf{E}_f^{(4)}$, $\mathbf{E}_f^{(8)}$ and $\mathbf{E}_f^{(11)}$ will excite eigenmodes $\mathbf{E}^{(5)}$, $\mathbf{E}^{(8)}$, $\mathbf{E}^{(4)}$, and $\mathbf{E}^{(8/4)}$,
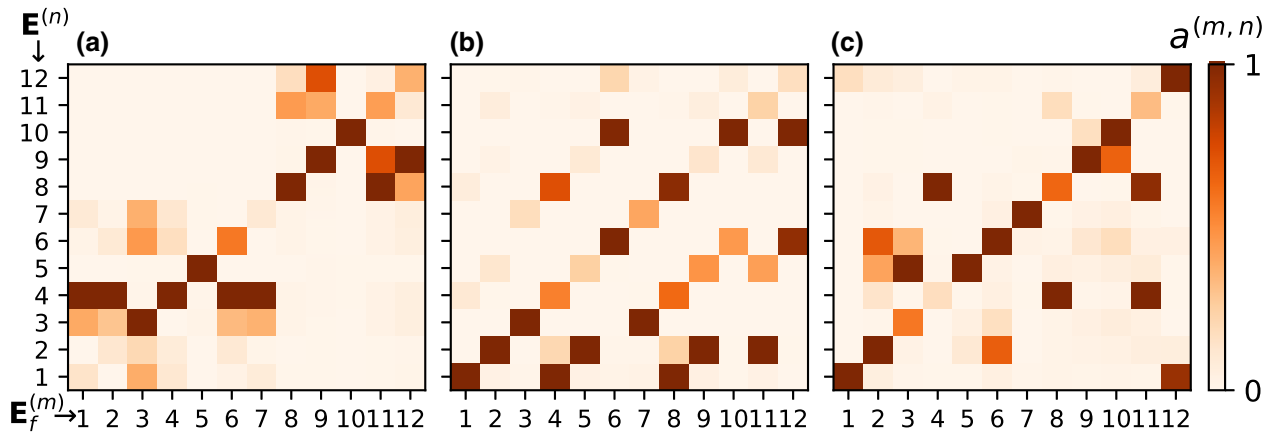
FIG. 9. The mode excitation coefficients between the designed focal fields $\mathbf{E}_f^{(m)}$ (along the horizontal axis) and eigenmodes $\mathbf{E}^{(n)}$ (along the vertical axis) for the nanodisk (a) monomer, (b) dimer, and (c) tetramer. For each designed focal field, the expansion coefficients for the first twelve eigenmodes are normalized to the maximum value.

respectively. The designed focal field $\mathbf{E}_f^{(12)}$ excites two eigenmodes simultaneously, including the one from which the focal field is inversely designed.

We note that some eigenmodes lead to inversely designed focal fields that cannot excite the eigenmode themselves. This is because a designed focal field is generated from the far-field components of the incident pupil field (the forward propagation) that is inversely designed from the far-field components of the eigenmode's near field (the backward propagation). During both the backward and forward propagation, high spatial harmonics (in Fourier space) are lost. As an intrinsic feature in any far-field approach, it is expected that a designed focal field might deviate from its original eigenmode near field. In addition, the necessary pupil field is only designed from the eigenmode's near field in a single plane (that coincides with the focal plane), and some necessary spatial harmonics might not be fully included. Therefore, it is reasonable that the inversely designed focal fields do not always excite the corresponding eigenmodes.

V. CONCLUSION AND DISCUSSION

We develop an inverse design approach of generating a focused light field that potentially matches with an eigenmode's near field in optical nanoantennas. We begin with a rigorous analysis of eigenmodes in the optical nanoantennas of interest. In this work, a nanodisk monomer, dimer, and tetramer are considered, and the first twelve eigenmodes in each nanostructure are studied using our BEM eigenmode solver that is based on the N-Müller formulation of the SIEs. In our inverse design strategies, the eigenmode's near field in the nanodisk's middle symmetric plane (which is chosen to coincide with the focal plane) is set as the desired focal field, which is inversely propagated to the pupil plane where the designed pupil field is obtained. In this backward propagation, we take into account the beamlike characteristic of the pupil field and the fact the pupil field is band limited, as well as the inherent dependencies of different field components. The eigenmodes are categorized into four groups by the time-averaged energies associated with the transverse and longitudinal electric and magnetic fields in the focal plane, and the dominant field components are chosen for uniquely determining the beamlike pupil field. The designed focal field's expansion coefficients into the first twelve eigenmodes are evaluated. For several eigenmodes, effective and exclusive mode excitations are achieved by designing the focal field accordingly from the eigenmodes via our developed inverse design method. This work can have a significant impact on optical switching, near- and far-field engineering in nano-optics.

[1] I. Staude, A. E. Miroshnichenko, M. Decker, N. T. Fofang, S. Liu, E. Gonzales, J. Dominguez, T. S. Luk, D. N. Neshev, I. Brener, and Y. Kivshar, Tailoring directional scattering through magnetic and electric resonances in subwavelength silicon nanodisks, ACS Nano **7**, 7824 (2013).

[2] K. E. Chong, I. Staude, A. James, J. Dominguez, S. Liu, S. Campione, G. S. Subramania, T. S. Luk, M. Decker, D. N. Neshev, I. Brener, and Y. S. Kivshar, Polarization-independent silicon metadevices for efficient optical wavefront control, Nano Lett. **15**, 5369 (2015).

[3] D. G. Baranov, R. Verre, P. Karpinski, and M. Käll, Anapole-enhanced intrinsic Raman scattering from silicon nanodisks, ACS Photonics **5**, 2730 (2018).

[4] X. Zhao and B. M. Reinhard, Switchable chiroptical hot-spots in silicon nanodisk dimers, ACS Photonics **6**, 1981 (2019).

[5] A. I. Kuznetsov, A. E. Miroshnichenko, M. L. Brongersma, Y. S. Kivshar, and B. Luk'yanchuk, Optically resonant dielectric nanostructures, Science **354**, aag2472 (2016).

[6] T. G. Habteyes, I. Staude, K. E. Chong, J. Dominguez, M. Decker, A. Miroshnichenko, Y. Kivshar, and I. Brener, Near-field mapping of optical modes on all-dielectric silicon nanodisks, ACS Photonics **1**, 794 (2014).

[7] J. Cambiasso, G. Grinblat, Y. Li, A. Rakovich, E. Cortés, and S. A. Maier, Bridging the gap between dielectric nanophotonics and the visible regime with effectively lossless GaP antennas, Nano Lett. **17**, 1219 (2017).

[8] R. Camacho-Morales, G. Bautista, X. Zang, L. Xu, L. Turquet, A. Miroshnichenko, H. H. Tan, A. Lamprianidis, M. Rahmani, C. Jagadish, D. N. Neshev, and M. Kauranen, Resonant harmonic generation in AlGaAs nanoantennas probed by cylindrical vector beams, Nanoscale **11**, 1745 (2019).

[9] J. D. Sautter, L. Xu, A. E. Miroshnichenko, M. Lysevych, I. Volkovskaya, D. A. Smirnova, R. Camacho-Morales, K. Zangeneh Kamali, F. Karouta, K. Vora, H. H. Tan, M. Kauranen, I. Staude, C. Jagadish, D. N. Neshev, and M. Rahmani, Tailoring second-harmonic emission from (111)-GaAs nanoantennas, Nano Lett. **19**, 3905 (2019).

[10] M. R. Shcherbakov, D. N. Neshev, B. Hopkins, A. S. Shorokhov, I. Staude, E. V. Melik-Gaykazyan, M. Decker, A. A. Ezhov, A. E. Miroshnichenko, I. Brener, A. A. Fedyanin, and Y. S. Kivshar, Enhanced third-harmonic generation in silicon nanoparticles driven by magnetic response, Nano Lett. **14**, 6488 (2014).

[11] S. A. Maier and H. A. Atwater, Plasmonics: Localization and guiding of electromagnetic energy in metal/dielectric structures, J. Appl. Phys. **98**, 011101 (2005).

[12] K. L. Kelly, E. Coronado, L. L. Zhao, and G. C. Schatz, The optical properties of metal nanoparticles: The influence of

size, shape, and dielectric environment, J. Phys. Chem. B **107**, 668 (2003).

[13] F. Moreno, P. Albella, and M. Nieto-Vesperinas, Analysis of the spectral behavior of localized plasmon resonances in the near-and far-field regimes, Langmuir **29**, 6715 (2013).

[14] J. Mäkitalo, M. Kauranen, and S. Suuriniemi, Modes and resonances of plasmonic scatterers, Phys. Rev. B **89**, 165429 (2014).

[15] Q. Zhan, Cylindrical vector beams: from mathematical concepts to applications, Adv. Opt. Photonics **1**, 1 (2009).

[16] C. Rosales-Guzmán, B. Ndagano, and A. Forbes, A review of complex vector light fields and their applications, J. Opt. **20**, 123001 (2018).

[17] J. Sancho-Parramon and S. Bosch, Dark modes and Fano resonances in plasmonic clusters excited by cylindrical vector beams, ACS Nano **6**, 8415 (2012).

[18] D. E. Gómez, Z. Q. Teo, M. Altissimo, T. J. Davis, S. Earl, and A. Roberts, The dark side of plasmonics, Nano Lett. **13**, 3722 (2013).

[19] A. Yanai, M. Grajower, G. M. Lerman, M. Hentschel, H. Giessen, and U. Levy, Near-and far-field properties of plasmonic oligomers under radially and azimuthally polarized light excitation, ACS Nano **8**, 4969 (2014).

[20] S. Reich, N. S. Mueller, and M. Bubula, Selection rules for structured light in nanooligomers and other nanosystems, ACS Photonics **7**, 1537 (2020).

[21] X. Zang, G. Bautista, L. Turquet, T. Setälä, M. Kauranen, and J. Turunen, Efficient hybrid-mode excitation in plasmonic nanoantennas by tightly focused higher-order vector beams, J. Opt. Soc. Am. B. **38**, 521 (2021).

[22] P. Lalanne, W. Yan, K. Vynck, C. Sauvan, and J.-P. Hugonin, Light interaction with photonic and plasmonic resonances, Laser Photon. Rev. **12**, 1700113 (2018).

[23] D. A. Powell, Resonant dynamics of arbitrarily shaped meta-atoms, Phys. Rev. B **90**, 075108 (2014).

[24] D. R. Fredkin and I. D. Mayergoyz, Resonant Behavior of Dielectric Objects (Electrostatic Resonances), Phys. Rev. Lett. **91**, 253902 (2003).

[25] I. D. Mayergoyz, D. R. Fredkin, and Z. Zhang, Electrostatic (plasmon) resonances in nanoparticles, Phys. Rev. B **72**, 155412 (2005).

[26] D. E. Gómez, K. C. Vernon, and T. J. Davis, Symmetry effects on the optical coupling between plasmonic nanoparticles with applications to hierarchical structures, Phys. Rev. B **81**, 075414 (2010).

[27] T. Weiss, N. A. Gippius, S. G. Tikhodeev, G. Granet, and H. Giessen, Derivation of plasmonic resonances in the Fourier modal method with adaptive spatial resolution and matched coordinates, J. Opt. Soc. Am. A **28**, 238 (2011).

[28] D. A. Bykov, E. A. Bezus, and L. L. Doskolovich, Use of aperiodic Fourier modal method for calculating complex-frequency eigenmodes of long-period photonic crystal slabs, Opt. Express **25**, 27298 (2017).

[29] D. A. Bykov and L. L. Doskolovich, Numerical methods for calculating poles of the scattering matrix with applications in grating theory, J. Lightwave Technol. **31**, 793 (2013).

[30] R. N. S. Suryadharma, M. Fruhnert, I. Fernandez-Corbaton, and C. Rockstuhl, Studying plasmonic resonance modes of hierarchical self-assembled meta-atoms based on their transfer matrix, Phys. Rev. B **96**, 045406 (2017).

[31] H. Guo, B. Oswald, and P. Arbenz, 3-dimensional eigenmodal analysis of plasmonic nanostructures, Opt. Express **20**, 5481 (2012).

[32] M. Paulus and O. J. F. Martin, Light propagation and scattering in stratified media: A Green's tensor approach, J. Opt. Soc. Am. A **18**, 854 (2001).

[33] G. D. Bernasconi, J. Butet, and O. J. Martin, Mode analysis of second-harmonic generation in plasmonic nanostructures, J. Opt. Soc. Am. B **33**, 768 (2016).

[34] D. A. Powell, Interference between the Modes of an All-Dielectric Meta-Atom, Phys. Rev. Appl. **7**, 034006 (2017).

[35] G. Bautista, M. J. Huttunen, J. Mäkitalo, J. M. Kontio, J. Simonen, and M. Kauranen, Second-harmonic generation imaging of metal nano-objects with cylindrical vector beams, Nano Lett. **12**, 3207 (2012).

[36] G. Bautista, J. Mäkitalo, Y. Chen, V. Dhaka, M. Grasso, L. Karvonen, H. Jiang, M. J. Huttunen, T. Huhtio, H. Lipsanen, and M. Kauranen, Second-harmonic generation imaging of semiconductor nanowires with focused vector beams, Nano Lett. **15**, 1564 (2015).

[37] R. Kant, Superresolution and increased depth of focus: An inverse problem of vector diffraction, J. Mod. Opt. **47**, 905 (2000).

[38] J. J. M. Braat, P. Dirksen, A. J. E. M. Janssen, S. van Haver, and A. S. van de Nes, Extended Nijboer–Zernike approach to aberration and birefringence retrieval in a high-numerical-aperture optical system, J. Opt. Soc. Am. A **22**, 2635 (2005).

[39] M. R. Foreman, S. S. Sherif, P. R. T. Munro, and P. Török, Inversion of the Debye-Wolf diffraction integral using an eigenfunction representation of the electric fields in the focal region, Opt. Express **16**, 4901 (2008).

[40] K. Jahn and N. Bokor, Solving the inverse problem of high numerical aperture focusing using vector Slepian harmonics and vector Slepian multipole fields, Opt. Commun. **288**, 13 (2013).

[41] Z. Chen, T. Zeng, and J. Ding, Reverse engineering approach to focus shaping, Opt. Lett. **41**, 1929 (2016).

[42] G. Rui, J. Chen, X. Wang, B. Gu, Y. Cui, and Q. Zhan, Synthesis of focused beam with controllable arbitrary homogeneous polarization using engineered vectorial optical fields, Opt. Express **24**, 23667 (2016).

[43] P. Yla-Oijala and M. Taskinen, Well-conditioned Muller formulation for electromagnetic scattering by dielectric objects, IEEE Trans. Antennas Propag. **53**, 3316 (2005).

[44] J. A. Stratton and L. J. Chu, Diffraction theory of electromagnetic waves, Phys. Rev. **56**, 99 (1939).

[45] S. A. Schelkunoff, Some equivalence theorems of electromagnetics and their application to radiation problems, Bell Syst. Tech. J. **15**, 92 (1936).

[46] W. C. Chew, M. S. Tong, and B. Hu, *Integral Equation Methods for Electromagnetic and Elastic Waves, Synthesis Lectures on Computational Electromagnetics* (Morgan & Claypool, San Rafael, Calif., 2009).

[47] S. J. Orfanidis, *Electromagnetic Waves and Antennas* (https://www.ece.rutgers.edu/ orfanidi/ewa/, 2016).

[48] J.-M. Jin, *Theory and Computation of Electromagnetic Fields* (John Wiley & Sons, Inc., 2015), 2nd ed.

[49] R. E. Collin, *Field Theory of Guided Waves* (John Wiley & Sons, Inc., 1991), 2nd ed.

[50] J. A. Kong, *Electromagnetic Wave Theory* (EMW Publishing, Cambridge, MA, 2008).

[51] D. Colton and R. Kress, *Integral Equation Methods in Scattering Theory* (Society for Industrial and Applied Mathematics, Philadelphia, PA, 2013).

[52] M. Taskinen, *On the Implementation and Formulation of the Electromagnetic Surface Integral Equations*, Ph.D. thesis, Helsinki University of Technology 2006.

[53] J. D. Jackson, *Classical Electrodynamics* (Wiley, New York, 1998), 3rd ed.

[54] E. Wolf, Electromagnetic diffraction in optical systems - I. An integral representation of the image field, Proc. R. Soc. Lond. A. Math. Phys. Sci. **253**, 349 (1959).

[55] B. Richards and E. Wolf, Electromagnetic diffraction in optical systems, II. Structure of the image field in an aplanatic system, Proc. R. Soc. Lond. A. Math. Phys. Sci. **253**, 358 (1959).

[56] S. F. Pereira and A. S. Van de Nes, Superresolution by means of polarisation, phase and amplitude pupil masks, Opt. Commun. **234**, 119 (2004).

[57] M. Leutenegger, R. Rao, R. Leitgeb, and T. Lasser, Fast focus field calculations, Opt. Express **14**, 11277 (2006).

[58] G.-L. Zhang, X.-Z. Gao, Y. Pan, M.-D. Zhao, D. Wang, H.-H. Zhang, Y. Li, C. Tu, and H.-T. Wang, Inverse method to engineer uniform-intensity focal fields with arbitrary shape, Opt. Express **26**, 16782 (2018).

[59] E. Wolf, A scalar representation of electromagnetic fields: II, Proc. Phys. Soc. **74**, 269 (1959).

[60] A. W. Snyder and J. D. Love, *Optical Waveguide Theory* (Chapman & Hall, London, 1983).

[61] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevApplied.18.044053 for a full list of scaled field components at different stages in the inverse design of all twelve eigenmodes in the nanodisk monomer, dimer, and tetramer.

[62] M. Frigo and S. G. Johnson, The design and implementation of FFTW3, Proc. IEEE **93**, 216 (2005).

[63] E. W. Cheney and D. R. Kincaid, *Numerical Mathematics and Computing* (Thomson Brooks/Cole, Belmont, CA, 2007), 6th ed.