

Advanced control of nonlinear beams with Pancharatnam-Berry metasurfacesM. Tymchenko,¹ J. S. Gomez-Diaz,² J. Lee,³ N. Nookala,¹ M. A. Belkin,¹ and A. Alù^{1,*}¹*Department of Electrical & Computer Engineering, The University of Texas at Austin, Austin, Texas 78712, USA*²*Department of Electrical of Computer Engineering, University of California, Davis, California 95616, USA*³*School of Electrical of Computer Engineering, Ulsan National Institute for Science and Technology, Ulsan 689-798, South Korea*

(Received 28 August 2016; revised manuscript received 12 November 2016; published 14 December 2016)

The application of the Pancharatnam-Berry (PB) phase approach to the design of nonlinear metasurfaces has recently enabled subdiffractive phase control over the generated nonlinear fields, embedding phased array features in ultrathin structures. Here, we rigorously model, analyze, and design highly efficient nonlinear metasurfaces with advanced functionalities, including the generation of pencil beams steered in arbitrary directions in space, as well as vortex beams with polarization-dependent angular momentum, and we extend the PB approach to various nonlinear processes. To this purpose, we develop an accurate and efficient theoretical framework—inspired by the linear phase array theory—based on the effective nonlinear susceptibility method, thus avoiding the use of time-consuming numerical simulations. Our findings allow exploiting the flat nonlinear optics paradigm, enabling exciting applications based on subwavelength field control over flat and large-scale structures with giant nonlinear responses.

DOI: [10.1103/PhysRevB.94.214303](https://doi.org/10.1103/PhysRevB.94.214303)**I. INTRODUCTION**

Nonlinear metasurfaces have recently provided record-high conversion efficiencies in nonlinear processes and hold a great potential to revolutionize the field of nonlinear optics by replacing bulk nonlinear crystals with flat structures of submicrometer thicknesses [1–4]. Strong nonlinear responses from such electrically small volumes require light-matter interactions much stronger than what is attainable in bulk crystals. This is where the field of plasmonics provides powerful tools. The use of carefully engineered subwavelength plasmonic inclusions offers a flexible and efficient way to engage strong fields in small volumes and to boost the efficiency of nonlinear processes, such as second-harmonic generation (SHG), to very large values [1,2,4–7]. In addition, ultrathin metasurfaces significantly alleviate phase matching constraints, which are of critical importance for efficient nonlinear processes [1,8]. Several attempts have been recently pursued to apply phase control techniques, which have been originally developed in linear optics, to nonlinear systems, aiming to provide much-needed control over generated fields at subwavelength scales [3,4,9–11]. Such nonlinear systems with wavefront engineering capabilities are paving the way towards a new paradigm in nonlinear optics, based on which advanced functionalities such as pencil beam steering, focusing, generation of vortex beams, holographic imaging, etc. are realized using ultrathin nonlinear metasurfaces, eliminating the need for bulky optical lenses and filters and mitigating the challenges associated with phase matching.

Recently, we introduced a novel platform for nonlinear metasurfaces, able to provide giant SHG efficiencies and simultaneously manipulate the emerging wavefront at will [9,12]. Specifically, we applied the Pancharatnam-Berry (PB) geometrical phase approach to nonlinear metasurfaces consisting of engineered split-ring plasmonic resonators loaded with nonlinear multi-quantum wells (MQWs). The basic functional-

ity of this approach consists in realizing metasurfaces formed by polarization-sensitive elements with spatially varying orientation. Under circularly polarized (CP) illumination, such elements generate local, nonlinear fields of equal magnitudes and controlled phases of CP components (see Appendix A for a detailed discussion) [13–15]. Similar techniques are used in linear optics to realize beam steering [13,16,17], focusing and defocusing of CP waves in reflection and transmission [15,16,18,19], polarization transformations [20], as well as to produce elaborated phase profiles, such as for Airy [21–23] and vortex beams [14,24–27]. In Refs. [9] and [12], we applied the geometrical phase approach to nonlinear metasurfaces, adiabatically rotating subwavelength plasmonic resonators in order to tailor right-handed and left-handed circularly polarized (RCP and LCP) second harmonic (SH) wavefront profiles. Due to the lack of efficient methods to model nonlinear systems of large size and complexity, our previous work applied this paradigm only to simple structures made of elements rotated following a one-dimensional (1D) phase gradient scheme. Yet, simulation of nonlinear metasurfaces composed of just a dozen of unit cells already requires tremendous computational resources and cannot be performed on compact desktop computers. In order to circumvent this issue, in Ref. [9] we introduced a semianalytical technique able to approximate the far field response of those nonlinear metasurfaces whose elements are rotated adiabatically along one direction, under the assumption of normally impinging pump beams.

In this contribution, we propose, design, and analyze two-dimensional (2D) nonlinear PB metasurfaces able to simultaneously provide high nonlinear conversion efficiency and advanced wave front shaping functionalities. Some examples of such functionalities include steering nonlinear pencil beams in arbitrary directions and the generation of vortex beams with different orbital angular momentum (OAM) for RCP and LCP components. We also extend these ideas to various nonlinear processes such as SHG, third harmonic generation (THG), and sum-frequency generation (SFG), and we describe how our wavefront tailoring approach can be applied to such

*alu@mail.utexas.edu

processes. To this purpose, we develop a general theoretical framework for modeling and predicting far-field radiation patterns of large-scale 2D nonlinear PB metasurfaces operating both in reflection and transmission and illuminated at arbitrary directions. This framework is inspired by the linear phased array theory for radio-frequency applications [28], but it is developed here in the realm of nonlinear metasurfaces with wavefront shaping capabilities implemented using the PB phase approach. Importantly, our technique relies on modeling the surface as an effective nonlinear susceptibility tensor, and therefore it can be applied to metasurfaces composed of any material undergoing arbitrary nonlinear processes. We do remark that generated fields significantly weaker than the pump field(s) and other common assumptions underlying in the development of our theory, as detailed below, are expected to be fulfilled in the common operation of these metasurfaces. In the following, we show how using this framework one may easily design and analyze 1D and 2D nonlinear metasurfaces, eliminating the need for extensive nonlinear numerical simulations and focusing instead on the design of unit cells with the highest possible nonlinear conversion efficiency.

The remainder of the paper is organized as follows. In Sec. II, we derive a general theoretical framework to characterize the far-field response of nonlinear PB metasurfaces. We show how this framework can be applied to metasurfaces aimed at second-, third-, and sum-frequency generation and discuss the assumptions and restrictions of our approach. In Sec. III, we rigorously validate our theory by comparing it against full-wave numerical results obtained using COMSOL [29]. For comparison purposes, we design and analyze a set of realistic SHG and THG nonlinear metasurfaces with 1D gradient based on highly efficient plasmonic resonators printed on MQWs. In addition, we determine under which conditions our proposed theory is accurate. Then, in Sec. IV we demonstrate advanced nonlinear PB metasurfaces with 2D gradients made of hundreds of elements and able to simultaneously provide high conversion efficiency and enhanced functionalities, such as shaping and steering nonlinear pencil beams in arbitrary directions and polarization-dependent vortex beam generation.

II. GENERAL THEORETICAL FRAMEWORK

In this section, we present a theoretical method to characterize the far-field response of ultrathin metasurfaces composed of subwavelength unit cells loaded with a nonlinear material. Specifically, we introduce a nonlinear phased array framework based on an effective nonlinear susceptibility model. In this approach, valid for conversion efficiencies below $\sim 5\text{--}10\%$ [1,30], the collective far-field response of the nonlinear PB metasurface is analytically predicted from the effective nonlinear susceptibility tensor of a single unit cell. This technique is first introduced for SHG and then extended to other nonlinear processes.

Our main goal is to design and analyze nonlinear PB metasurfaces able to manipulate the wavefront of the generated signal at will. For this purpose, the structure must fulfill several requirements imposed by the PB phase approach [9] that we briefly list here. First, in order to provide phase control of the generated nonlinear field, the unit cells must be substantially

smaller than the wavelength of the pump *and* nonlinear signals of interest. This imposes certain restrictions on the size and spacing between neighboring elements. Second, each cell must contain a specifically designed polarization-sensitive PB phase element (PB element) that responds only to certain field polarizations at the fundamental or generated harmonic frequency, or both. Third, the orientation of PB elements must change adiabatically so that each element is surrounded by alike neighbors. Fourth, these PB elements must be designed in such a way that the cross-coupling between neighboring elements is minimized so that they can be rotated independently from each other without largely affecting their individual response. Unlike phased arrays operating in the linear regime, the cross-coupling between nonlinear PB elements may introduce large phase and amplitude corrections that are difficult to model analytically. However, if the coupling between neighboring unit cells is weak, such metasurfaces become an ideal platform for tailoring the nonlinear wavefront in a straightforward fashion.

We begin our analysis by considering a 2D array of unit cells located at $\mathbf{r}_n = (x_n, y_n, 0)$, where n is the index number. Under the assumptions given above, we can describe their second-order nonlinear response with an effective nonlinear susceptibility tensor $\overset{\leftrightarrow}{\chi}^{(2)}(\mathbf{r}_n; 2\omega : \omega, \omega)$ (see Appendix B and Refs. [1] and [30] for a detailed procedure), which relates the nonlinear polarization density $\mathbf{P}_n^{2\omega}$ induced in the n th unit cell averaged over its volume and oscillating at the SH frequency 2ω with the impinging (pump) plane wave \mathbf{E}^ω oscillating at the fundamental frequency (FF) ω :

$$\mathbf{P}_n^{2\omega} = \varepsilon_0 \overset{\leftrightarrow}{\chi}^{(2)}(\mathbf{r}_n; 2\omega : \omega, \omega) : \mathbf{E}^\omega(\mathbf{r}_n)\mathbf{E}^\omega(\mathbf{r}_n), \quad (1)$$

where the colon denotes a double-dot dyadic product. We stress that the effective susceptibility encapsulates all linear scattering effects and relates the average induced nonlinear polarization density in every cell to the *incident* field. This technique assumes the generation of weak nonlinear fields (compared to the pump), thus avoiding the need for solving the electromagnetic problem self-consistently at both pump and generated frequencies. This is a valid assumption, as typical efficiency levels achieved in the most efficient nonlinear metasurfaces reported to date are below a few percent.

Following this approach, first we compute the effective nonlinear susceptibility tensor $\overset{\leftrightarrow}{\chi}^{(2)}$ in Cartesian coordinates assuming close to normal incidence. Taking advantage of the weak cross-coupling between adjacent cells and the fact that they differ from each other only by the PB elements' orientations ψ_n (see Fig. 1), the local effective susceptibility tensor of the n th cell can be obtained as

$$\overset{\leftrightarrow}{\chi}^{(2)}(\mathbf{r}_n) \equiv \overset{\leftrightarrow}{\chi}^{(2)}(\psi_n) = \mathbf{R}(\psi_n) \cdot \overset{\leftrightarrow}{\chi}^{(2)} : \mathbf{R}(-\psi_n)\mathbf{R}(-\psi_n), \quad (2)$$

where \mathbf{R} is a rotation matrix around the z axis (see Fig. 1),

$$\mathbf{R}(\psi) = \begin{pmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (3)$$

The SH electric field $\mathbf{E}^{2\omega}(\mathbf{r})$ generated by the entire metasurface can be found as a sum of radiation from independent effective dipole moments of each cell, $\mathbf{d}_n^{2\omega} = V_{\text{UC}}\mathbf{P}_n^{2\omega}$ with

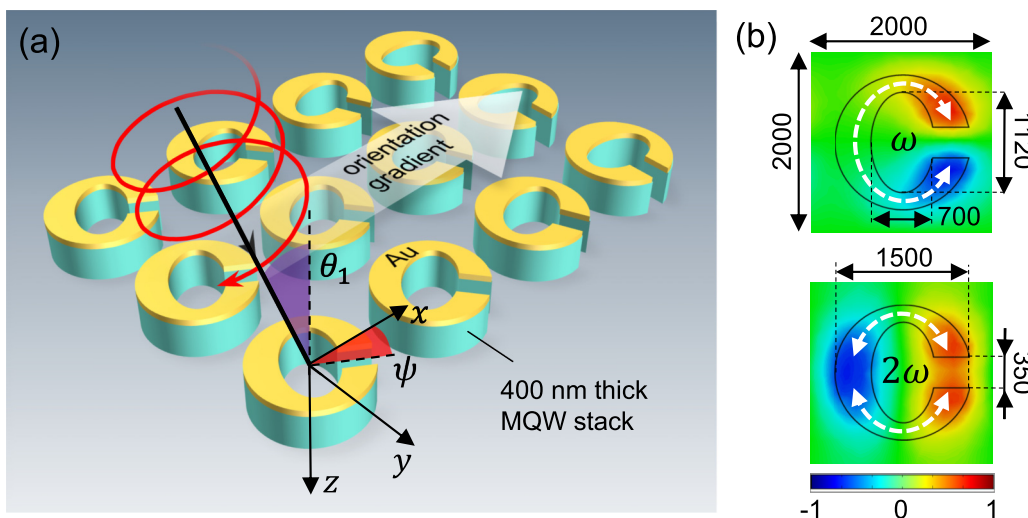


FIG. 1. (a) Schematic of a nonlinear metasurface consisting of PB optical elements designed for efficient SHG. Each element contains a nonlinear material, in this case a 400-nm-thick MQW semiconductor heterostructure stack with semiconductor layers in xy plane with gold plasmonic resonators placed on top. The MQW is etched around the resonator in order to minimize the cross-coupling between the cells. The orientation of optical elements, ψ , varies linearly along the x -axis. The metasurface is illuminated at an incident angle θ_1 ($\varphi_1 = 0$). (b) E_z field distribution for a reference (not rotated) unit cell at the fundamental and SH frequency, ω and 2ω , respectively. The dimensions of the unit cell are specified in nm. The shape of the resonator is chosen so that it responds to y -polarized field at ω and x polarization at 2ω , at the same time ensuring a subwavelength square footprint.

V_{UC} denoting the unit-cell volume,

$$\mathbf{E}^{2\omega}(\mathbf{r}) = \frac{k_2^2}{\epsilon_0} \sum_n \vec{G}(\mathbf{r}, \mathbf{r}_n) \cdot \mathbf{d}_n^{2\omega}, \quad (4)$$

where $k_2 = 2\omega/c$, with c denoting the speed of light in free space and $\vec{G}(\mathbf{r}, \mathbf{r}_n)$ is a dyadic Green's function that takes into account the influence of the media surrounding the metasurface. Without loss of generality, we assume here that the metasurfaces under study are suspended in free space, allowing us to employ the well-known far-field free-space Green's function,

$$\vec{G}_{\text{FF}}(\mathbf{r}, \mathbf{r}_n) = [\vec{I} - \hat{\mathbf{r}}\hat{\mathbf{r}}] \frac{e^{ik_2 r}}{4\pi r} e^{-ik_2 \hat{\mathbf{r}} \cdot \mathbf{r}_n}, \quad (5)$$

with $\mathbf{r} = r\hat{\mathbf{r}}$ [31]. More complex scenarios, including the presence of substrates and ground planes, can easily be modeled considering a modified Green's function.

To easily account for the physical rotation of different PB elements, it is convenient to write Eqs. (1)–(5) in CP basis. Let $|a_u\rangle = |R_u, L_u, r_u\rangle$ denote a CP polarization state corresponding to RCP, LCP, and longitudinal (radial) polarization components of incoming and radiated waves, where the subscript u is the index of a wave (for the SHG case, $u = 1$

corresponds to the input pump wave and $u = 2$ denotes the output SH wave). Each basis can be defined uniquely by a pair of polar angles (θ_u, φ_u) in spherical coordinates that correspond to the propagation direction of the wave, with $\theta_u \in [0, \pi]$ and $\varphi_u \in [0, 2\pi]$,

$$|R_u\rangle = \frac{1}{\sqrt{2}}(\hat{\theta}_u - i\hat{\varphi}_u), \quad |L_u\rangle = \frac{1}{\sqrt{2}}(\hat{\theta}_u + i\hat{\varphi}_u), \\ |z_u\rangle = \hat{\mathbf{r}}_u, \quad (6)$$

where $\hat{\theta}_u \equiv \hat{\theta}(\theta_u, \varphi_u)$, $\hat{\varphi}_u \equiv \hat{\varphi}(\theta_u, \varphi_u)$, $\hat{\mathbf{r}}_u \equiv \hat{\mathbf{r}}(\theta_u, \varphi_u)$ are basis unit vectors of the spherical coordinate system. We also introduce a circular polarization basis $|a_0\rangle \equiv |a\rangle = |R, L, z\rangle$ associated with the metasurface itself and corresponding to CP waves propagating perpendicularly to the metasurface in the $+z$ direction ($\theta = \varphi = 0$), see Fig. 1(a):

$$|R\rangle = \frac{1}{\sqrt{2}}(\hat{\mathbf{x}} - i\hat{\mathbf{y}}), \quad |L\rangle = \frac{1}{\sqrt{2}}(\hat{\mathbf{x}} + i\hat{\mathbf{y}}), \quad |z\rangle = \hat{\mathbf{z}}. \quad (7)$$

All CP bases, $|a\rangle$ and $|a_u\rangle$, can be related to a Cartesian polarization basis $|i\rangle = |x, y, z\rangle$ through the coordinate transformation $|a_u\rangle = \Lambda_u|i\rangle$, where $\Lambda_u = \Lambda(\theta_u, \varphi_u)$ is a unitary coordinate transformation matrix given by

$$\Lambda(\theta, \varphi) = \frac{1}{\sqrt{2}} \begin{pmatrix} \cos\theta \cos\varphi + i \sin\varphi & \cos\theta \sin\varphi - i \cos\varphi & -\sin\theta \\ \cos\theta \cos\varphi - i \sin\varphi & \cos\theta \sin\varphi + i \cos\varphi & -\sin\theta \\ \sqrt{2} \sin\theta \cos\varphi & \sqrt{2} \sin\theta \sin\varphi & \sqrt{2} \cos\theta \end{pmatrix}. \quad (8)$$

In the CP basis $|a\rangle$, the nonlinear susceptibility tensor elements can be found by performing the coordinate transformation

$$\chi_{abc}^{(2)} = \sum_{ijk} [\Lambda_0]_{ai} \chi_{ijk}^{(2)} [\Lambda_0^{-1}]_{jb} [\Lambda_0^{-1}]_{kc}, \quad (9)$$

with $i, j, k = \{x, y, z\}$, $a, b, c = \{R, L, z\}$, and

$$\mathbf{\Lambda}_0 \equiv \mathbf{\Lambda}_{\left. \begin{array}{l} \theta=0 \\ \psi=0 \end{array} \right|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i & 0 \\ 1 & i & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix}. \quad (10)$$

Using a similar transformation, the rotation matrix in the CP basis $|a\rangle$ can be found from (3) as

$$R_{ab} = \sum_{ij} [\mathbf{\Lambda}_0]_{ai} R_{ij} [\mathbf{\Lambda}_0^{-1}]_{jb} = \begin{pmatrix} e^{-i\psi} & 0 & 0 \\ 0 & e^{i\psi} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (11)$$

Since the matrix is diagonal, by substituting (11) into (2) it is easy to see that

$$\chi_{abc}^{(2)}(\psi_n) = \chi_{abc}^{(2)} e^{i\psi_n(a-b-c)}, \quad (12)$$

where each of the indexes a , b , and c in the exponential factor should be replaced according to (11) as follows: $R = -1$, $L = +1$, and $z = 0$.

The phase factor $e^{i\psi_n(a-b-c)}$ is only due to the local element orientation, and thus it is of a purely geometrical nature. The emergence of this geometrical phase can be intuitively understood by noting that the adiabatic rotation of adjacent resonators by an angle ψ provides [9] the following: (i) an extra phase factor of $e^{\mp i\psi}$ to the RCP/LCP components of

the surface currents induced in the metasurface and (ii) an additional factor $e^{\pm i\psi}$ to the RCP/LCP components of the beam reflected in the structure. Despite its apparent simplicity, the emergence of this phase is a manifestation of the celebrated PB geometrical phase (see Appendix A), which is of profound importance in optics and quantum mechanics. In metasurfaces, geometrical phase gradients break the inversion symmetry $\mathbf{r} \rightarrow -\mathbf{r}$ by imprinting a transverse momentum $\nabla\psi(\mathbf{r})$ that leads to a splitting of the dispersion relation of states with opposite optical helicity [32]. This effect is similar to Rashba spin-band splitting in 2D electronic systems subjected to a transverse potential gradient [33,34].

Using (6), an arbitrary polarized plane wave \mathbf{E}^ω with wavenumber $\mathbf{k}_1 = k_1 \hat{\mathbf{r}}_1$, where $k_1 = \omega/c$ (c is the speed of light in free space), obliquely impinging onto a metasurface, can be described as

$$\mathbf{E}^\omega(\mathbf{r}_n) = \sum_{a_1} |a_1\rangle E_{a_1}^\omega e^{i\mathbf{k}_1 \cdot \mathbf{r}}. \quad (13)$$

Projecting it onto the basis $|a\rangle$, we obtain the corresponding amplitudes

$$\mathbf{E}^\omega(\mathbf{r}) = \sum_a |a\rangle \langle a| \sum_{a_1} |a_1\rangle E_{a_1}^\omega e^{i\mathbf{k}_1 \cdot \mathbf{r}} = \sum_a |a\rangle E_a^\omega e^{i\mathbf{k}_1 \cdot \mathbf{r}}, \quad (14)$$

where $E_a^\omega = \sum_{a_1} \langle a | a_1 \rangle E_{a_1}^\omega$ and $\langle a | = |a\rangle^\dagger$ is the standard bra-ket notation used in quantum mechanics, with the dagger denoting the Hermitian adjoint. The inner product $\langle a | a_1 \rangle$ is given as (see Appendix C)

$$\langle a | a_1 \rangle = \frac{1}{2} \begin{bmatrix} (\cos \theta_1 + 1)e^{i\varphi_1} & (\cos \theta_1 - 1)e^{i\varphi_1} & \sqrt{2} \sin \theta_1 e^{i\varphi_1} \\ (\cos \theta_1 - 1)e^{-i\varphi_1} & (\cos \theta_1 + 1)e^{-i\varphi_1} & \sqrt{2} \sin \theta_1 e^{-i\varphi_1} \\ -\sqrt{2} \sin \theta_1 & -\sqrt{2} \sin \theta_1 & 2 \cos \theta_1 \end{bmatrix}_{a, a_1}. \quad (15)$$

In the same basis, $|a\rangle$, the dipole moment is $\mathbf{d}_n^{2\omega} = \sum_a |a\rangle d_{n,a}^{2\omega}$ with

$$d_{n,a}^{2\omega} = V_{\text{UC}} \sum_{bc} \chi_{abc}^{(2)} E_b^\omega E_c^\omega e^{i[2\mathbf{k}_1 \cdot \mathbf{r} + \psi_n(a-b-c)]}. \quad (16)$$

Performing a projection onto the CP basis $|a_2\rangle$ associated with the observation direction, we find

$$d_{n,a_2}^{2\omega} = V_{\text{UC}} \sum_{abc} \langle a_2 | a \rangle \chi_{abc}^{(2)} E_b^\omega E_c^\omega e^{i[2\mathbf{k}_1 \cdot \mathbf{r} + \psi_n(a-b-c)]}, \quad (17)$$

where

$$\langle a_2 | a \rangle = \frac{1}{2} \begin{bmatrix} (\cos \theta_2 + 1)e^{-i\varphi_2} & (\cos \theta_2 - 1)e^{i\varphi_2} & -\sqrt{2} \sin \theta_2 \\ (\cos \theta_2 - 1)e^{-i\varphi_2} & (\cos \theta_2 + 1)e^{i\varphi_2} & -\sqrt{2} \sin \theta_2 \\ \sqrt{2} \sin \theta_2 e^{-i\varphi_2} & \sqrt{2} \sin \theta_2 e^{i\varphi_2} & 2 \cos \theta_2 \end{bmatrix}_{a_2, a} \quad (18)$$

(see Appendix C). We want to stress that, differently from [9] where we defined the RCP and LCP components of the polarization density with respect to $\pm z$ directions (transmission and reflection regions, respectively), here all CP components in $|a\rangle$ are defined uniquely with respect to the $+z$ direction. Thus, due to the coordinate transformation (18), if $\theta_2 = 0$ (propagation along $+z$), we have $d_{n,R_2}^{2\omega} = d_{n,R}^{2\omega}$, but if $\theta_2 = \pi$ (propagation along $-z$), we obtain $d_{n,L_2}^{2\omega} = d_{n,R}^{2\omega}$, i.e., the RCP nonlinear polarization density becomes a source for the LCP component of SH radiation in the reflection region. Combining Eqs. (14)–(18) we obtain

$$\mathbf{d}_n^{2\omega} = \sum_{a_2} |a_2\rangle d_{n,a_2}^{2\omega}, \quad (19)$$

with

$$d_{n,a_2}^{2\omega} = \varepsilon_0 V_{\text{UC}} \sum_{abc} \langle a_2 | a \rangle \chi_{abc}^{(2)} A_b^\omega A_c^\omega e^{i2\mathbf{k}_1 \cdot \mathbf{r} + i\psi_n(a-b-c)}, \quad (20)$$

where all summations are performed over all three polarization components. In addition, it can be easily shown that in the CP basis $|a_2\rangle$, the free-space dyadic Green's function (5) can be written as

$$\vec{G}_{\text{FF}}(\mathbf{r}, \mathbf{r}_n) = \frac{e^{ik_2 r}}{4\pi r} e^{-ik_2 \hat{\mathbf{f}}_2 \cdot \mathbf{r}_n} \begin{pmatrix} |R_2\rangle\langle R_2| - |r_2\rangle\langle r_2| & & \\ & |L_2\rangle\langle L_2| - |r_2\rangle\langle r_2| & \\ & & 0 \end{pmatrix}. \quad (21)$$

Substituting (19)–(21) into (4) and taking into account that $\langle a_2 | b_2 \rangle = 0$ if $a_2 \neq b_2$, we finally obtain

$$\mathbf{E}^{2\omega}(\mathbf{r}) = \sum_{a_2} |a_2\rangle E_{a_2}^{2\omega}(\mathbf{r}), \quad (22)$$

$$E_{a_2}^{2\omega}(\mathbf{r}) = V_{\text{UC}} k_2^2 \frac{e^{ik_2 r}}{4\pi r} \sum_{a b c} \langle a_2 | a \rangle \chi_{abc}^{(2)} \text{AF}_{abc}^{\text{SHG}} E_b^\omega E_c^\omega, \quad (23)$$

where the index $a_2 = \{R_2, L_2\}$ and the third (radial) component is strictly zero. The tensor $\text{AF}_{abc}^{\text{SHG}}$ is a nonlinear array factor for the SHG process, given as

$$\text{AF}_{abc}^{\text{SHG}} = \sum_n e^{i[(2\mathbf{k}_1 - k_2 \hat{\mathbf{f}}_2) \cdot \mathbf{r}_n + \psi_n(a - b - c)]}. \quad (24)$$

From Eqs. (23) and (24), it is evident that the SH radiation in the far field is composed of the sum of all array factors weighted by the corresponding nonlinear susceptibility tensor element in the CP basis.

This result can be generalized to other nonlinear processes. In particular, for THG processes in which the generated field oscillates at frequency 3ω , we can write

$$E_{a_2}^{3\omega}(\mathbf{r}) = V_{\text{UC}} \frac{k_2^2}{4\pi r} e^{ik_2 r} \sum_{a b c} \langle a_2 | a \rangle \chi_{abcd}^{(3)} E_b^\omega E_c^\omega E_d^\omega \text{AF}_{abcd}^{\text{THG}}, \quad (25)$$

with $a_2 = \{R_2, L_2\}$, $k_2 = 3\omega/c$ and a THG array factor

$$\text{AF}_{abcd}^{\text{THG}} = \sum_n e^{i[(3\mathbf{k}_1 - k_2 \hat{\mathbf{f}}_2) \cdot \mathbf{r}_n + \psi_n(a - b - c - d)]}. \quad (26)$$

For SFG with two pump waves \mathbf{E}^{ω_1} and \mathbf{E}^{ω_2} oscillating at ω_1 and ω_2 , respectively, the radiated wave will oscillate at frequency $\omega_3 = \omega_1 + \omega_2$,

$$E_{a_3}^{\omega_3}(\mathbf{r}) = V_{\text{UC}} \frac{k_3^2}{4\pi r} e^{ik_3 r} \sum_{a b c} \langle a_3 | a \rangle \chi_{abc}^{(2)} E_b^{\omega_1} E_c^{\omega_2} \text{AF}_{abc}^{\text{SFG}}, \quad (27)$$

with $a_3 = \{R_3, L_3\}$, $k_3 = \omega_3/c$, $E_b^{\omega_1} = \sum_{b_1} \langle b | b_1 \rangle E_{b_1}^{\omega_1}$, $E_c^{\omega_2} = \sum_{c_2} \langle c | c_2 \rangle E_{c_2}^{\omega_2}$, and

$$\text{AF}_{abc}^{\text{SFG}} = \sum_n e^{i[(\mathbf{k}_1 + \mathbf{k}_2 - k_3 \hat{\mathbf{f}}_3) \cdot \mathbf{r}_n + \psi_n(a - b - c)]}. \quad (28)$$

This formulation allows fast computation of the field radiated by ultrathin nonlinear metasurfaces, and it explicitly accounts for the fact that the pump signals can impinge obliquely. In addition to the restrictions already imposed by the PB phase approach on the nonlinear metasurface design, this framework relies on the accurate evaluation of the effective nonlinear susceptibility tensor of a single unit cell in a periodic environment. This tensor encapsulates all scattering effects at the fundamental and generated frequencies, and

thus its components also depend on the illumination and observation angle, showing significantly different values in endfire/backfire directions compared to broadside. Here, we imply that the effective nonlinear susceptibility tensor has been numerically evaluated, assuming close-to-normal incidence and radiation directions. As a result, the accuracy of our theoretical analysis is expected to decrease for incident and generated beams propagating at large angles with respect to the normal to the metasurface, especially if the unit cell is not too small compared to the wavelength.

III. THE 1D GRADIENT NONLINEAR METASURFACES: THEORY AND SIMULATIONS

In this section, we validate the accuracy of the formulation derived in the previous section by performing a direct comparison with numerical simulations. To this purpose, we present, study, and discuss specific nonlinear metasurfaces composed of PB elements with a linear orientation gradient along one direction, able to provide 1D beam-scanning functionalities.

We consider a host nonlinear metasurface aimed at highly efficient SHG process [9]: an array of gold split-ring resonators (SRRs) placed on top of MQW substrates etched around the SRRs in order to reduce the cross-coupling [see Fig. 1(a)]. Here we assume that the elements are suspended in free-space, thus allowing them to radiate both in reflection and transmission. We also assume that MQWs provide a large second-order intrinsic response at the pump frequency of 30 THz. The SRR dimensions are accordingly chosen to provide resonant field enhancement at FF 30 THz and SH frequency 60 THz [see Fig. 1(b)]. It is important to mention that MQWs present a nonlinear response only for electric field components orthogonal to the semiconductor layers [35–41], i.e., they respond only to z -polarized components of FF fields and generate z -polarized nonlinear polarization SH currents. Doubly-resonant SRRs provide a large overlap integral (as discussed in Appendix B and in Refs. [1] and [7]) between the z components of FF and SH fields [see Fig. 1(b)], ensuring a SH nonlinear response of the metasurface many orders of magnitude larger than the one typically obtained in bulk nonlinear crystals [1,4]. As a result, SHG conversion efficiencies of nearly 0.1% were demonstrated experimentally in mid-infrared spectral range ($\lambda = 3..15 \mu\text{m}$) using metasurfaces only 400–600 nm thick [4,7]. Moreover, the conversion efficiency can be boosted to values above 4% using high-quality MQWs heterostructures and optimized plasmonic resonators, as discussed in Refs. [7], [28], and [30].

Full-wave numerical simulations of nonlinear metasurfaces were performed using COMSOL [29]. Specifically, each structure was simulated in frequency domain in two steps: First, the entire metasurface consisting of $N = 24$ unit cells along the x direction and infinitely periodic along y was tested

at ω ; then, the nonlinear polarization currents driven by the local field at ω were impressed in MQW volumes, and the structure was simulated at the generated frequency. As long as the nonlinear field remains significantly weaker than the pump, this numerical approach is rigorous and applicable to nonlinear metasurfaces with any desired functionality. In practice, such numerical simulations take a long time and require large computational resources, thus limiting their applicability to the analysis of metasurfaces composed of even a few unit cells, hindering the fast design of realistic structures. A full-wave numerical analysis of one of the nonlinear metasurfaces described below in this section requires many computational hours of a powerful dedicated workstation.

On the contrary, the framework developed in the previous section can be efficiently applied to the analysis and design of these nonlinear metasurfaces, while providing physical insight into the metasurface operation. For example, in the specific case of 1D nonlinear metasurfaces with a linear orientation gradient only along the x direction, we can particularize and further simplify the proposed theoretical formalism. Specifically, consider a 1D array of N unit cells with coordinates $\mathbf{r}_n = \hat{x}nd$, where d is the length of the unit cell along x . A linear gradient allows us to write the local PB element orientation as $\psi_n = n\Delta\psi$, where $\Delta\psi$ is the orientation variation step along x . Assuming the incident wave impinging in the xz plane, i.e., $\mathbf{k}_1 = k_1(\sin\theta_1, 0, \cos\theta_1)$, from Eq. (24) we obtain

$$\text{AF}_{abc}^{\text{SHG}} = \sum_{n=1}^N e^{in[(2k_1 \sin\theta_1 - k_2 \sin\theta_2 \cos\varphi_2)d + \Delta\psi(a-b-c)]}. \quad (29)$$

Introducing the notation

$$X_m = (2k_1 \sin\theta_1 - k_2 \sin\theta_2 \cos\varphi_2)d + m\Delta\psi, \quad (30)$$

the summation in (29) can be performed explicitly, leading to an expression which is similar to the array factor in the linear regime,

$$\text{AF}_{abc}^{\text{SHG}} = e^{iX_{a-b-c}(N-1)/2} \left[\frac{\sin\left(\frac{N}{2}X_{a-b-c}\right)}{\sin\left(\frac{1}{2}X_{a-b-c}\right)} \right]. \quad (31)$$

Unlike the case of linear phased arrays, here we cannot drop the exponential phase prefactor because the far-field expression (23) contains a summation over all indexes of the tensor and we should keep the accurate phase difference between all terms. From (31), it follows that in a nonlinear PB metasurface with a linear orientation gradient, all components of the nonlinear susceptibility tensor $\chi_{abc}^{(2)}$ contribute to the radiation in specific directions defined by maxima of Eq. (31). The direction of the main lobes in reflection and transmission domains from each term, θ_m^R and θ_m^T , can be found from the condition $X_m = 0$, which yields

$$\theta_m^T = \pm \sin^{-1} \left(\sin\theta_1 + m \frac{\Delta\psi}{2\pi} \frac{\lambda_{2\omega}}{d} \right), \quad (32)$$

$$\theta_m^R = \pi - \theta_m^T. \quad (33)$$

The choice between ‘+’ and ‘-’ corresponding to $\varphi_2 = 0$ and $\varphi_2 = \pi$, respectively, is made so that θ_m^T is positive. Inspecting Eq. (32), we conclude that in the transmission

region a normally incident RCP pump wave at ω produces a RCP SH wave ($m = +1$) at 2ω steered closer to $+z$ direction than the LCP component ($m = +3$). In reflection, the situation is reversed, and the RCP wave will be steered at a larger angle with the $-z$ direction; in turn, the LCP wave will propagate closer to the broadside.

Figure 2 shows beam steering capabilities of SHG metasurfaces, computed by the proposed theory and numerical simulations. Figures 2(a) and 2(b) show the SH far-field response with orientation step $\Delta\psi = 15^\circ$ for normally impinging RCP and LCP beams, respectively. Our results confirm the expected separation of directions for RCP and LCP components of the SH signal [9]. Importantly, an excellent agreement between theoretically and numerically computed directivities is obtained. According to Eq. (32), the two main lobes for RCP and LCP components in transmission occur at $\theta_{+1}^T \approx 7^\circ$ and $\theta_{+3}^T \approx 18^\circ$, respectively, and in the reflection region the main lobes occur at $\theta_{+3}^R \approx 162^\circ$ for RCP and $\theta_{+1}^R \approx 173^\circ$ for LCP component [see arrows in Fig. 2(a)]. We emphasize that these directions can be controlled nearly arbitrarily (though, not independently from each other) following the simple equations provided above. Figures 2(c) and 2(d) show similar results, but considering now a nonlinear metasurface with the orientation step $\Delta\psi = 30^\circ$. As expected, the RCP and LCP generated beams are now directed towards a larger angle with respect to the z axis. One can see that the discrepancies between numerical results and theory away from broadside slightly increase. As discussed in Sec. II, such differences arise because the effective susceptibility tensor is evaluated at normal incidence in a periodic environment, and it becomes less accurate for larger gradient step. Nevertheless, it is seen that the theoretical directivities still accurately predict directions and magnitudes of the two main lobes both in transmission and reflection.

We would like to emphasize that the proposed nonlinear metasurface provides beam steering capabilities also when illuminated by the pump at oblique incidence. A tangential momentum of the impinging wave is added to the momentum produced by the phase gradient imprinted in the metasurface, enabling continuous SHG beam-scanning functionality. In Fig. 3, we show theoretical and numerical results for the RCP and LCP polarized Gaussian beams impinging in the direction $\theta_1 = 20^\circ$, $\varphi_1 = 0$, on a PB metasurface with $\Delta\psi = 15^\circ$. For RCP incidence [Fig. 3(a)], the momentum from the gradient is added to the longitudinal momentum of the impinging beam, increasing the steering angles. In turn, for a LCP beam, the imprinted phase variation tends to compensate for the longitudinal impinging wave’s momentum and to restore a steering direction closer to the broadside.

Finally, we show that the proposed concept can also be applied to other nonlinear processes, such as THG. In practice, constructing a THG metasurface with phase control is a challenging task that requires designing a plasmonic PB element able to provide sufficiently strong field enhancement at ω and 3ω in a subwavelength footprint, which goes beyond the focus of this paper. However, since here we are primarily interested in far-field directivity, we can take advantage of the fact that the unit-cell design used for SHG is polarization sensitive at ω and that the phase of the FF field in MQW still depends on the SRRs’ orientation. Additionally, even at the

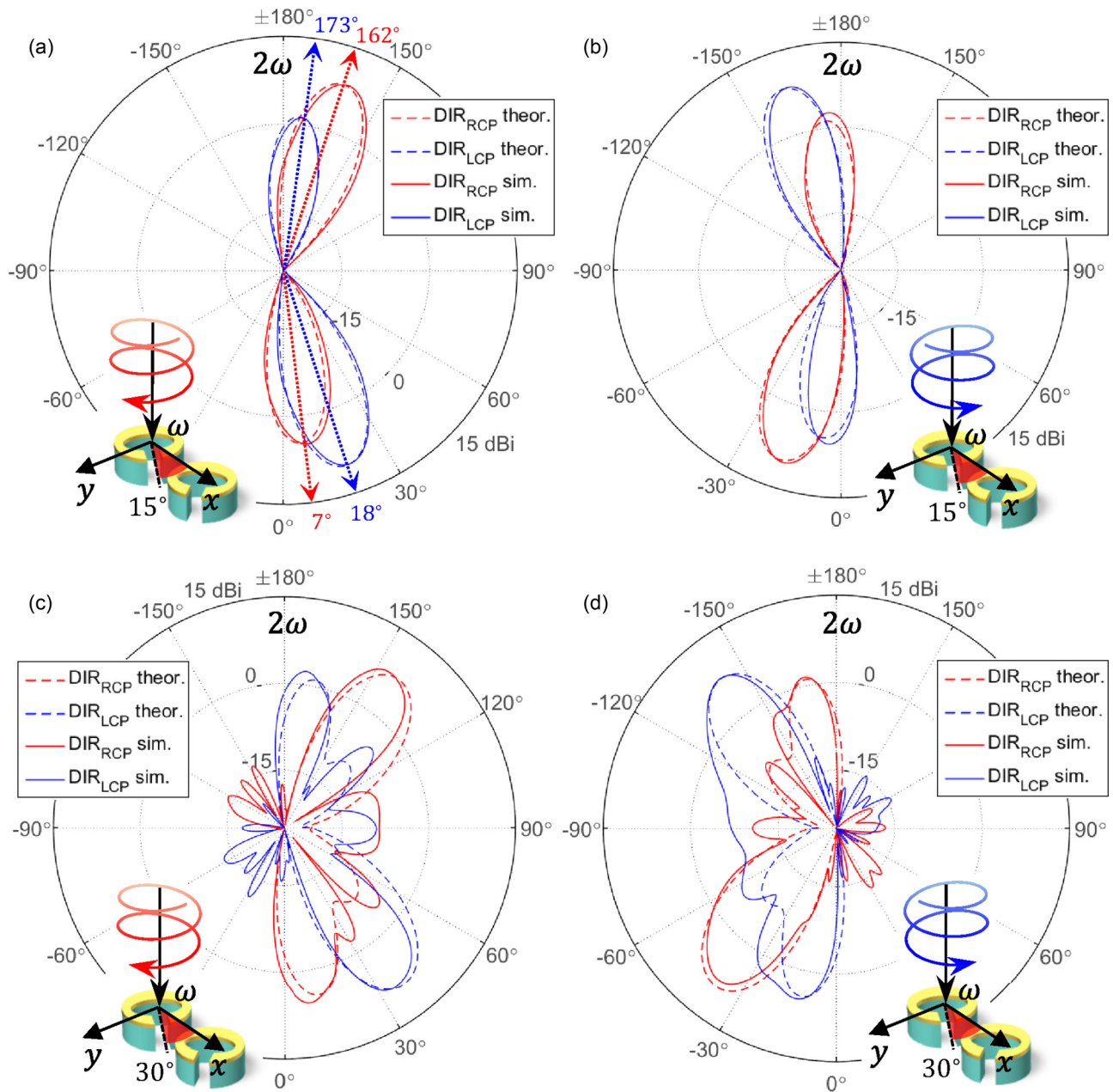


FIG. 2. Far-field directivity patterns (in dBi) for RCP and LCP components of SH radiation generated by a nonlinear metasurface illuminated by a Gaussian beam of $12 \mu\text{m}$ full width at half maximum (FWHM). The directivity plots are shown in the xz plane. Theoretical results are shown with dashed lines, and solid lines show results from full-wave simulations. (a) A RCP polarized beam is impinging at normal incidence, $\theta_1 = 0$ on the metasurface with the rotation step $\Delta\psi = 15^\circ$ (see the inset). The dotted lines with arrow tips show theoretically predicted highest partial directivities. (b) LCP incidence, $\theta_1 = 0$, $\Delta\psi = 15^\circ$. (c), (d) The same as (a) and (b) but for $\Delta\psi = 30^\circ$.

angular frequency 3ω , some polarization selectivity is present thanks to the asymmetrical shape of the resonator and the fact that periodicity is still smaller than the third-harmonic wavelength ($d/\lambda_{3\omega} \approx 0.6$). Consequently, the metasurface designed for SHG is also able to provide THG (90 THz for a 30 THz pump) but with much lower conversion efficiency. Figure 4 confirms THG and beam-steering capabilities for a nonlinear metasurface with $\Delta\psi = 15^\circ$ under normal incidence, computed by our proposed theory and validated using full-wave simulations in COMSOL. Our results confirm (i) the

presence of THG and the capability to manipulate the radiated beam and (ii) the accuracy of the proposed theory to model metasurfaces employing various nonlinear phenomena. The discrepancy between analytical and numerical results comes from the fact that the unit cell is larger than half wavelength at the third harmonic so that the effective susceptibility model loses part of its accuracy. Still, the proposed theory provides a reasonably accurate general picture of the two main CP lobes of the THG signal steered away from the normal.

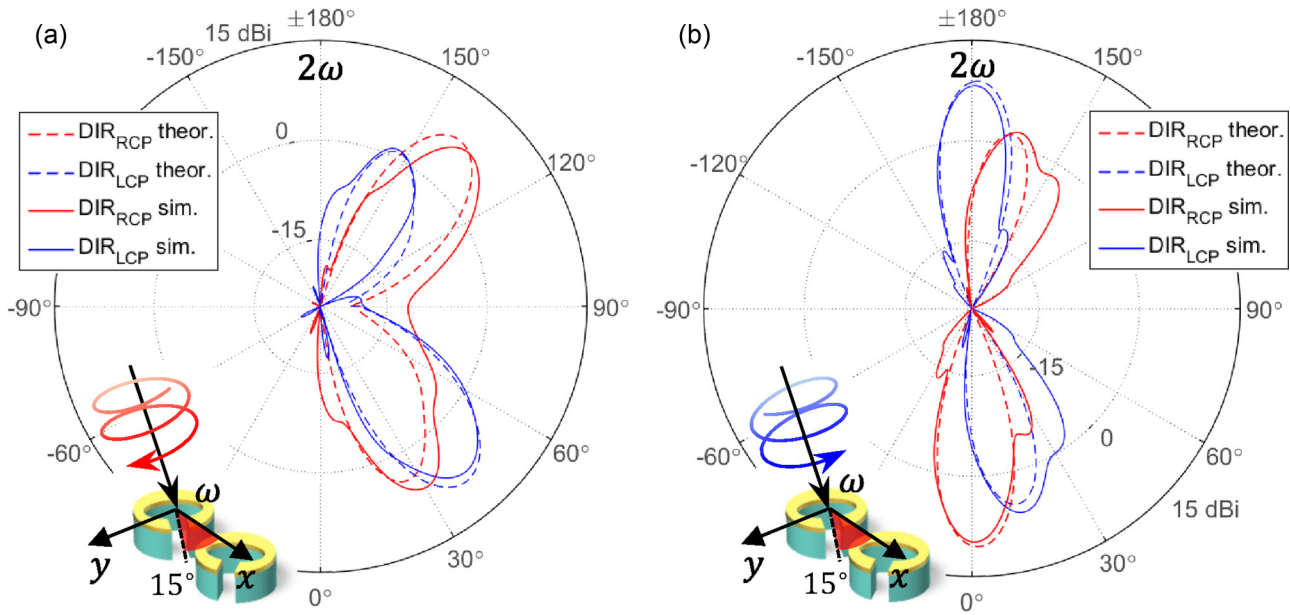


FIG. 3. Far-field partial directivity patterns (in dBi) for RCP and LCP components of SH radiation from the metasurface orientation step $\Delta\psi = 15$ but at oblique incidence, $\theta_1 = 20^\circ$. Directivity plots are shown in the xz plane. (a) RCP incident beam; (b) LCP incident beam.

IV. METASURFACES WITH 2D PHASE PATTERNS: ADVANCED FUNCTIONALITIES

In this section, we apply the theoretical framework developed in previous sections to propose and demonstrate advanced 2D nonlinear metasurfaces with interesting functionalities, such as the generation of pencil beams directed towards arbitrary directions in space and vortex beams with different angular momentum for RCP and LCP polarizations. We stress that we were unable to perform the simulation of such metasurfaces using commercial full-wave software due to the

required amount of computational resources, so we limit our results to designs and calculations based on our analytical model validated in the previous section.

The first example that we consider is a SH nonlinear metasurface with a 2D gradient profile, as illustrated in Fig. 5(a), providing beam-steering capabilities. In general, beam steering in arbitrary directions can be achieved by adjusting the orientation steps along the x and y directions, $\Delta\psi_x$ and $\Delta\psi_y$, respectively. Following the procedure introduced in the previous section, the 2D array factor of this

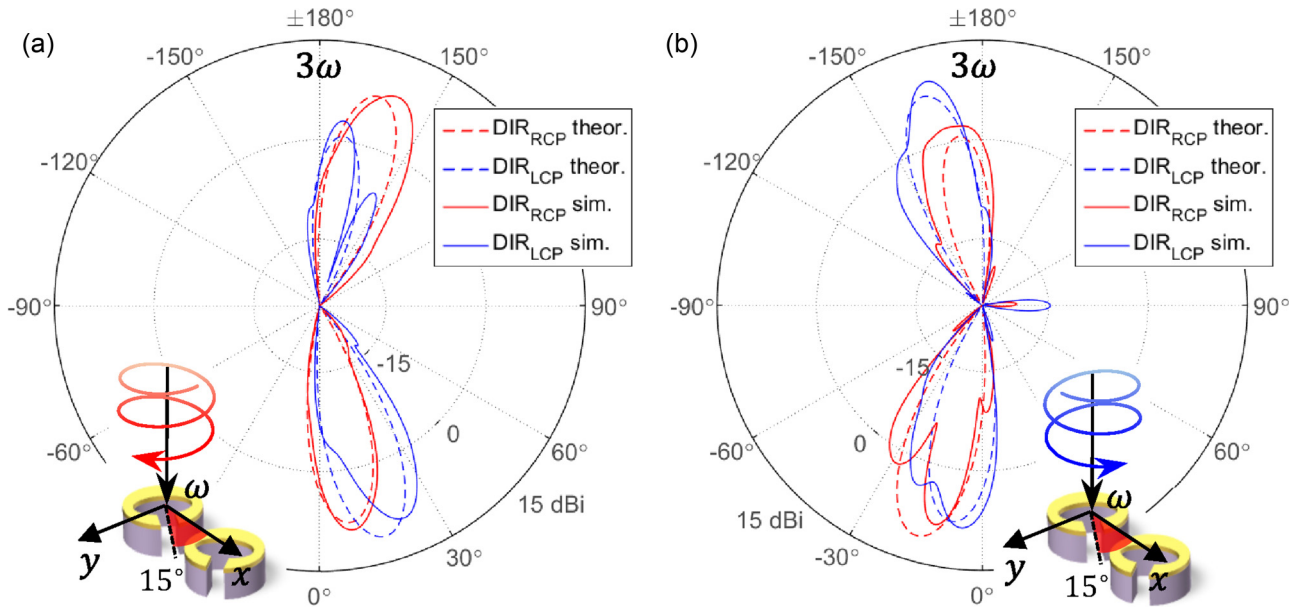


FIG. 4. Far-field partial directivity patterns (in dBi) for third-harmonic radiation from the same metasurface, as in Figs. 2(a) and 2(b), with third-order nonlinearity described by an effective nonlinear susceptibility tensor $\vec{\chi}^{(3)}$. (a) The RCP pump Gaussian beam at the frequency ω impinging normally; (b) LCP pump beam.

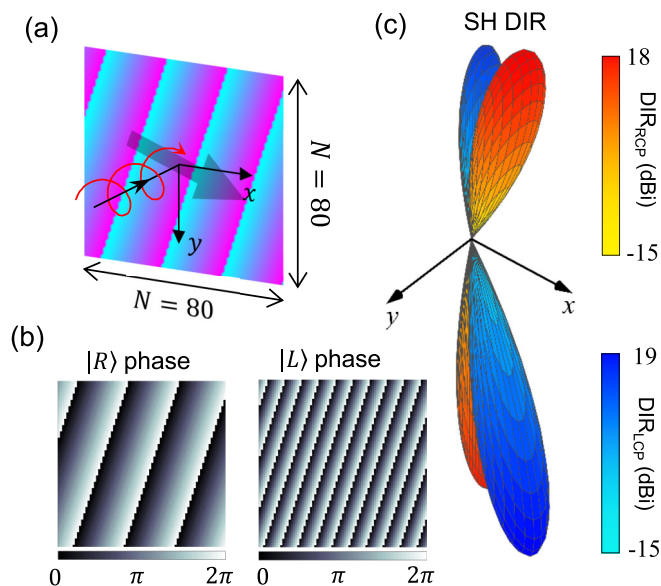


FIG. 5. (a) Schematic of a 80×80 cells SHG metasurface with a linear variation of elements' orientation in x - y direction (shown with an arrow). The metasurface is illuminated by a RCP Gaussian beam with FWHM of $23 \mu\text{m}$ at normal incidence. (b) Spatial phase profiles of RCP and LCP polarized components of the SH field near $z = 0$. (c) Three-dimensional (3D) partial directivity patterns for RCP and LCP polarized components of SH radiation.

metasurface is

$$\text{AF}_{abc}^{\text{SHG}} = e^{iX_{a-b-c}(N_x-1)/2} e^{iY_{a-b-c}(N_y-1)/2} \times \left[\frac{\sin\left(\frac{N_x}{2} X_{a-b-c}\right)}{\sin\left(\frac{1}{2} X_{a-b-c}\right)} \right] \left[\frac{\sin\left(\frac{N_y}{2} Y_{a-b-c}\right)}{\sin\left(\frac{1}{2} Y_{a-b-c}\right)} \right], \quad (34)$$

where N_x , N_y are the number of unit cells along the x and y direction, and

$$X_m = (2k_1 \sin \theta_1 \cos \phi_1 - k_2 \sin \theta_2 \cos \phi_2)d + m \Delta \psi_x, \quad (35)$$

$$Y_m = (2k_1 \sin \theta_1 \sin \phi_1 - k_2 \sin \theta_2 \sin \phi_2)d + m \Delta \psi_y. \quad (36)$$

The direction of the main lobes for each polarization combination in transmission, (θ_m^T, ϕ_m) , and reflection, (θ_m^R, ϕ_m) , can be approximately found from Eqs. (34)–(36) by simultaneously setting $X_m = Y_m = 0$, which yields

$$\theta_m^T = \sin^{-1} \left[\left(\sin \theta_1 \cos \phi_1 + m \frac{\Delta \psi_x}{2\pi} \frac{\lambda_{2\omega}}{d} \right)^2 + \left(\sin \theta_1 \sin \phi_1 + m \frac{\Delta \psi_y}{2\pi} \frac{\lambda_{2\omega}}{d} \right)^2 \right]^{1/2}, \quad (37)$$

$$\theta_m^R = \pi - \theta_m^T, \quad (38)$$

$$\phi_m = \tan^{-1} \left(\frac{\sin \theta_1 \sin \phi_1 + m \frac{\Delta \psi_y}{2\pi} \frac{\lambda_{2\omega}}{d}}{\sin \theta_1 \cos \phi_1 + m \frac{\Delta \psi_x}{2\pi} \frac{\lambda_{2\omega}}{d}} \right). \quad (39)$$

In Fig. 5, we show the results for a metasurface with linear gradient steps along x and y directions of $\Delta \psi_x = 15^\circ$ and $\Delta \psi_y = 5^\circ$, respectively. The phase profiles for the LCP and

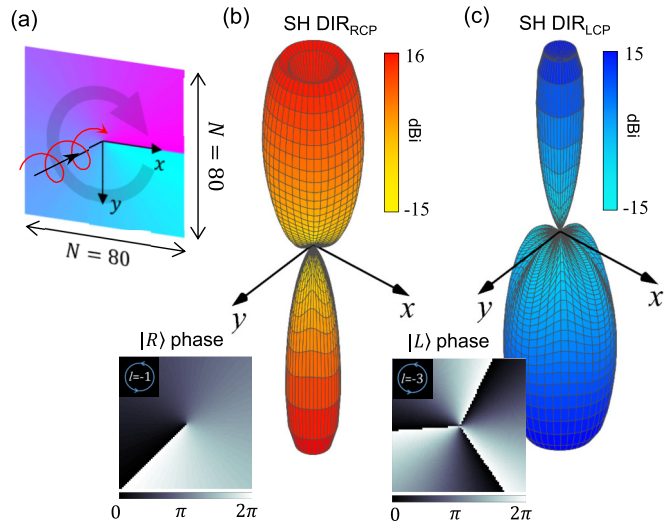


FIG. 6. (a) Schematic of a 80×80 SHG metasurface with a helical variation of elements' orientation $\psi(x, y) = \varphi(x, y)$ (shown with a purple arrow), where φ is the polar angle. The metasurface is illuminated by a RCP Gaussian beam with FWHM of $30 \mu\text{m}$ at normal incidence. Partial directivity patterns for RCP and LCP polarized SH components are shown in (b) and (c), respectively. The insets show spatial phase profiles of transmitted beams at $z = 0$. The RCP beam carries OAM $l = -1$, and for the LCP beam it is $l = -3$.

RCP components of the SH field in each cell at $z = 0$ are depicted in Fig. 5(b). Figure 5(c) shows theoretically computed partial directivity patterns for the RCP and LCP components of the SH radiation. Results confirm a high degree of control over the generated beams even for small gradients and the possibility of continuous steering of the RCP and LCP beams towards almost any direction that is not too far away from the broadside.

In the second example, we show a nonlinear PB metasurface generating SH helical beams with nonzero OAM, l . Differently from the linear case, where only one of the CP components of the transmitted or reflected SH field acquires geometrical phase, in case of SHG, both CP components are subject to geometrical phase patterning. As a result, using a fixed element's orientation pattern in a nonlinear metasurface, we obtain two oppositely polarized helical beams with different OAM. Figure 6(a) illustrates a SHG metasurface with spatial orientation dependence, $\psi(x, y) = \varphi(x, y)$, where $\varphi(x, y)$ is an azimuthal angle in the xy plane. Figures 6(b) and 6(c) show the RCP and LCP polarization components of the SH far field, clearly showing a doughnut shape cross-section profile (the field is instantaneous). The RCP component possesses $l = -1$, and for the LCP component, $l = -3$. The radius of the doughnut increases for larger $|l|$, as expected in this type of beam [42]. The corresponding phase profiles of RCP and LCP components at $z = 0$ are shown in the insets of Figs. 6(b) and 6(c).

Finally, we would like to demonstrate that the same metasurface can also work at oblique incidence, effectively steering the SH helical beams in different directions away from normal. In Fig. 7, we show the same metasurface with orientation pattern $\psi(x, y) = \varphi(x, y)$ but illuminated by an obliquely impinging beam ($\theta_1 = 15^\circ, \phi_1 = 30^\circ$). The resulting

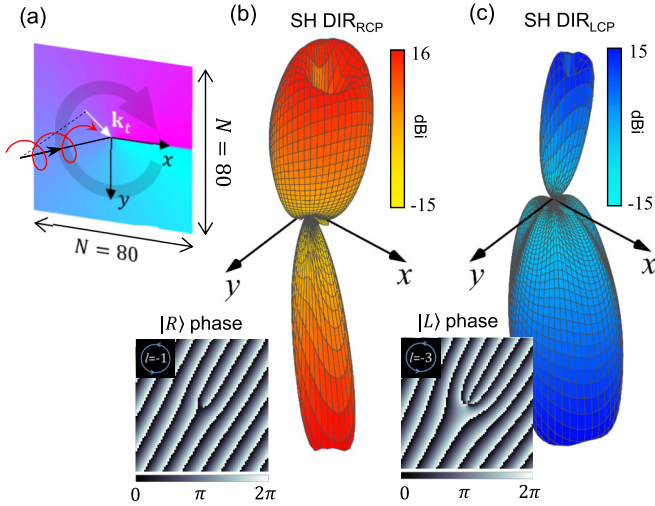


FIG. 7. The same metasurface with orientation profile $\psi(x, y) = \varphi(x, y)$, as in Fig. 6, but illuminated by an obliquely incident incidence. (a) Schematic of the setup. (b), (c) Partial directivities for RCP and LCP SH radiation, respectively. The insets show the corresponding phase profiles for transmitted beams.

longitudinal momentum $\mathbf{k}_t = k_1 \sin \theta_1 (\cos \varphi_1, \sin \varphi_1, 0)$ adds up to the momentum generated by the phase gradient, resulting in a well-known fork-shaped phase profile [10,43]. Note that this phase profile can be readily imprinted in the metasurface itself. If such metasurface is illuminated by a linearly polarized wave, the SH radiation will be shaped into helical beams, and opposite polarizations will be steered in the opposite side of the z axis. Similar responses have been successfully observed in the past using the quasiperiodic polling technique [10,43], which enforces substantial constraints on the size of the beam because the illuminated spot necessarily has to cover a large number of polling periods. In contrast, our approach provides a continuous phase shaping from ultrathin nonlinear metasurfaces with giant conversion efficiencies.

V. CONCLUSION

In this paper, we have presented a rigorous and efficient theoretical framework to characterize the far-field response of large-scale ultrathin nonlinear metasurfaces with phase shaping functionalities, implemented using the PB approach. Our formulation, validated using detailed comparison with results from full-wave commercial software, can be applied to various nonlinear processes such as SHG and THG and permits the rapid analysis and design of nonlinear PB metasurfaces with advanced functionalities. Specifically, we have demonstrated the concept of highly efficient nonlinear metasurfaces based on plasmonic resonators printed on top of MQWs, able to provide exciting features such as the generation of pencil-beam directed to any desired direction in space and vortex beams with polarization-dependent angular momentum. Our approach can be easily applied to the design of large-scale, highly efficient and advanced metasurfaces—composed of any nonlinear material—able to manipulate the generated beams at will.

ACKNOWLEDGMENTS

This paper was supported by the Air Force Office of Scientific Research, the Simons Foundation and the Welch foundation with Grant No. F-1802.

APPENDIX A: PB GEOMETRICAL PHASE

In classical optics, the PB geometrical phase is usually introduced for a beam whose polarization state undergoes a cyclic sequence of polarization transformations (i.e., the state is returned back to the initial one) that yields an extra phase factor in addition to a dynamical phase factor associated with the beam propagation [44,45]. Specifically, it has been shown that if the polarization state history of a beam when plotted on the Poincaré sphere closes a loop, the geometrical phase will be equal to a half of the geodesic area encompassed by this loop. A linear PB element performs a single transformation of an incident beam's polarization, and thus no geometrical phase is induced in the copolarized component because the encompassed geodesic area is zero. However, it can be shown that the geometric phase emerges in the polarization component orthogonal to the polarization state of the impinging beam, so the resulting state of the beam will be different at various orientations of the PB element. For example, for a RCP wave impinging on two similar PB elements at different orientations, the RCP component of the transmitted wave will have no geometrical phase difference, while the LCP component will develop a phase difference. In linear optics, this result has been extensively used to tailor the phase of the cross-polarized component of the transmitted beam [13–15,19,26,27,32]. As we demonstrated in Ref. [9], in the nonlinear case, both CP components of the generated field can be subject to geometrical phase change, thus allowing an extra degree of freedom.

APPENDIX B: EFFECTIVE NONLINEAR SUSCEPTIBILITY TENSOR RETRIEVAL

Consider an arbitrary polarized plane wave $\mathbf{E}_{\text{inc}}^\omega = E_{\text{inc},x}^\omega \hat{x} + E_{\text{inc},y}^\omega \hat{y}$ oscillating at frequency ω and impinging normally from one side of a thin metasurface consisting from identical unit cells, which have a local intrinsic second-order susceptibility tensor $\overleftrightarrow{\chi}^{(2)}(\mathbf{r})$ that can vary from point to point across the volume of each cell. Since the polarization components of the impinging wave are orthogonal to each other, each of them will independently induce its own portion of the local electric field inside a unit cell, $\mathbf{E}_{\text{UC}}^\omega(\mathbf{r}) = \mathbf{E}_{\text{UC}(x)}^\omega(\mathbf{r}) + \mathbf{E}_{\text{UC}(y)}^\omega(\mathbf{r})$, each of which can be found from full-wave simulations. Here we focus on SHG, so that in what follows we imply $\overleftrightarrow{\chi}^{(2)}(\mathbf{r}) \equiv \overleftrightarrow{\chi}^{(2)}(\mathbf{r}, 2\omega : \omega, \omega)$, and later we explain how this method can be extended to other nonlinear processes. For SHG, the total local second-order polarization density oscillating at 2ω can be found as

$$\mathbf{P}_{\text{UC}}^{2\omega}(\mathbf{r}) = \varepsilon_0 \overleftrightarrow{\chi}^{(2)}(\mathbf{r}) : \mathbf{E}_{\text{UC}}^\omega(\mathbf{r}) \mathbf{E}_{\text{UC}}^\omega(\mathbf{r}), \quad \mathbf{r} \in V_{\text{UC}}, \quad (\text{B1})$$

where V_{UC} denotes the volume of the unit cell. Since the metasurface is infinite and the unit cells are identical, in the far field it will emit a plane wave $\mathbf{E}_{\text{FF}}^{2\omega}$ of yet unknown amplitude sustained by a nonlinear polarization current density

$\mathbf{J}_{\text{UC}}^{2\omega} = -i 2\omega \mathbf{P}_{\text{UC}}^{2\omega}$. In order to obtain the amplitude of this wave, we can apply Lorentz reciprocity theorem to relate $\mathbf{E}_{\text{FF}}^{2\omega}$ radiated by $\mathbf{J}_{\text{UC}}^{2\omega}$ and the field inside the unit cell $\mathbf{E}_{\text{UC}}^{2\omega}$ sustained by an imaginary uniform current in the far field, $\mathbf{K}_{\text{FF}}^{2\omega}$, emitting a plane wave $\mathbf{E}_{\text{inc}}^{2\omega}$. Similar to the field at ω , we can separate the portions excited by the x - and y -polarized incident field components, $\mathbf{E}_{\text{FF}(x)}^{2\omega}$ and $\mathbf{E}_{\text{FF}(y)}^{2\omega}$, respectively. Applying the reciprocity theorem, we obtain,

$$\int_{S_{\text{UC}}} \mathbf{E}_{\text{FF}}^{2\omega} \cdot \mathbf{K}_{\text{FF}}^{2\omega} dS = -i 2\omega \int_{V_{\text{UC}}} \mathbf{E}_{\text{UC}}^{2\omega}(\mathbf{r}) \cdot \mathbf{P}_{\text{UC}}^{2\omega}(\mathbf{r}) dV, \quad (\text{B2})$$

where S_{UC} is the area of the unit cell. From quasistatics, we find that $\mathbf{K}_{\text{FF}}^{2\omega} = 2\eta_0^{-1} \mathbf{E}_{\text{inc}}^{2\omega}$, where $\eta_0 = \sqrt{\mu_0/\epsilon_0}$ is a free-space impedance. Substituting Eq. (B1) into (B2) and performing the integration in the right-hand side (r.h.s.), we obtain

$$\mathbf{E}_{\text{inc}}^{2\omega} \cdot \mathbf{E}_{\text{FF}}^{2\omega} = \frac{-i\omega c^{-1}}{S_{\text{UC}}} \int_{V_{\text{UC}}} \mathbf{E}_{\text{UC}}^{2\omega}(\mathbf{r}) \cdot \overset{\leftrightarrow}{\chi}^{(2)}(\mathbf{r}) : \mathbf{E}_{\text{UC}}^{2\omega}(\mathbf{r}) \mathbf{E}_{\text{UC}}^{2\omega}(\mathbf{r}) dV, \quad (\text{B3})$$

where c is the speed of light in free space. Now we argue that in the quasistatic approximation, instead of $\mathbf{J}_{\text{UC}}^{2\omega}$, radiated SH field $\mathbf{E}_{\text{FF}}^{2\omega}$ can be sustained by an effective nonlinear polarization current $\mathbf{J}_{\text{eff}}^{2\omega} = -i 2\omega \mathbf{P}_{\text{eff}}^{2\omega}$. It is up to our choice how we define this effective nonlinear susceptibility density $\mathbf{P}_{\text{eff}}^{2\omega}$. We assume it is uniform across a thin layer of the same thickness as the metasurface, h_{cell} ; however, we would like to avoid solving the homogeneous scattering problem, so we assume that it has $\overset{\leftrightarrow}{\chi}^{(1)} = 0$, i.e., the layer is transparent at ω , and the whole scattering phenomena are now contained within $\overset{\leftrightarrow}{\chi}_{\text{eff}}^{(2)}$. In this case, the problem is greatly simplified, and we may now simply write $\mathbf{P}_{\text{eff}}^{2\omega} = \epsilon_0 \overset{\leftrightarrow}{\chi}_{\text{eff}}^{(2)} : \mathbf{E}_{\text{inc}}^{\omega} \mathbf{E}_{\text{inc}}^{\omega}$, and consequently

$$\mathbf{E}_{\text{FF}}^{2\omega} = -i\omega c^{-1} h_{\text{UC}} \overset{\leftrightarrow}{\chi}_{\text{eff}}^{(2)} : \mathbf{E}_{\text{inc}}^{\omega} \mathbf{E}_{\text{inc}}^{\omega}, \quad (\text{B4})$$

where h_{UC} is the unit-cell height. Substituting Eq. (B4) into (B3) we obtain

$$\begin{aligned} \mathbf{E}_{\text{inc}}^{2\omega} \cdot \overset{\leftrightarrow}{\chi}_{\text{eff}}^{(2)} : \mathbf{E}_{\text{inc}}^{\omega} \mathbf{E}_{\text{inc}}^{\omega} \\ = \frac{1}{V_{\text{UC}}} \int_{V_{\text{UC}}} \mathbf{E}_{\text{UC}}^{2\omega}(\mathbf{r}) \cdot \overset{\leftrightarrow}{\chi}^{(2)}(\mathbf{r}) : \mathbf{E}_{\text{UC}}^{\omega}(\mathbf{r}) \mathbf{E}_{\text{UC}}^{\omega}(\mathbf{r}) dV. \end{aligned} \quad (\text{B5})$$

Expanding the left-hand side of Eq. (B5) in x - and y -polarization components and separating the corresponding fields in the r.h.s., we obtain,

$$\begin{aligned} E_{\text{inc},i}^{2\omega} \sum_{jk} \chi_{\text{eff},ijk}^{(2)} E_{\text{inc},j}^{\omega} E_{\text{inc},k}^{\omega} \\ = \frac{1}{V_{\text{UC}}} \int_{V_{\text{UC}}} \mathbf{E}_{\text{UC}(i)}^{2\omega}(\mathbf{r}) \cdot \overset{\leftrightarrow}{\chi}^{(2)}(\mathbf{r}) : \\ \times \sum_{jk} \mathbf{E}_{\text{UC}(j)}^{\omega}(\mathbf{r}) \mathbf{E}_{\text{UC}(k)}^{\omega}(\mathbf{r}) dV, \end{aligned} \quad (\text{B6})$$

where $i, j, k = \{x, y\}$. Finally, equating the terms with the same indexes, we obtain an elegant yet very powerful expression

$$\begin{aligned} \chi_{\text{eff},ijk}^{(2)} = \frac{1}{V_{\text{UC}}} \int_{V_{\text{UC}}} \left[\frac{\mathbf{E}_{\text{UC}(i)}^{\omega}(\mathbf{r})}{E_{\text{inc},i}^{\omega}} \right] \cdot \overset{\leftrightarrow}{\chi}^{(2)}(\mathbf{r}) : \\ \times \left[\frac{\mathbf{E}_{\text{UC}(j)}^{\omega}(\mathbf{r})}{E_{\text{inc},j}^{\omega}} \right] \left[\frac{\mathbf{E}_{\text{UC}(k)}^{\omega}(\mathbf{r})}{E_{\text{inc},k}^{\omega}} \right] dV. \end{aligned} \quad (\text{B7})$$

From Eq. (B7), it is evident that the each component of the nonlinear susceptibility tensor is equal to an overlap integral between the fields induced in a unit cell by an i -polarized plane wave at 2ω and a jk -polarization combination of the field at ω , weighted by a local value of the intrinsic susceptibility of the nonlinear medium and averaged over the unit-cell volume. By altering the overlap integral, we can engineer virtually any component $\chi_{\text{eff},ijk}^{(2)}$ from any $\chi_{ijk}^{(2)}$. For instance, for a MQW-loaded unit cell with only $\chi_{zzz}^{(2)}$ nonzero tensor element, we find that

$$\begin{aligned} \chi_{\text{eff},xyz}^{(2)} = \frac{1}{V_{\text{UC}}} \int_{V_{\text{UC}}} \chi_{zzz}^{(2)}(\mathbf{r}) \left[\frac{E_{\text{UC}(x),z}^{2\omega}(\mathbf{r})}{E_{\text{inc},x}^{2\omega}} \right] \\ \times \left[\frac{E_{\text{UC}(y),z}^{\omega}(\mathbf{r})}{E_{\text{inc},y}^{\omega}} \right] \left[\frac{E_{\text{UC}(z),z}^{\omega}(\mathbf{r})}{E_{\text{inc},z}^{\omega}} \right] dV. \end{aligned} \quad (\text{B8})$$

The indexes of the effective susceptibility tensor's elements do not span over the z polarization. This is a consequence of the fact that in the derivation of (B7), we assumed normal incidence and radiation. In order to evaluate all the components of the $\overset{\leftrightarrow}{\chi}_{\text{eff}}^{(2)}$ tensor, we need nonzero $E_{\text{inc},z}^{\omega}$ and $E_{\text{inc},z}^{2\omega}$, which can be achieved by allowing the impinging and radiated waves to propagate at small angles with respect to the z axis and subtracting the field contribution of x and y components of the incident field. Finally, Eq. (B8) holds not only when the pump and generated signal come from the same side of the metasurface but also for any particular configuration of incidence and radiation beams, as long as they are not too far from the z axis.

APPENDIX C: TRANSFORMATION MATRICES CALCULATION

The inner product $\langle a | a_1 \rangle$ appearing in transformations of amplitudes from the polarization basis $|a_1\rangle$ to the basis $|a\rangle$ can be found by presenting each of these states through a Cartesian basis $|i\rangle$ with $i = \{x, y, z\}$,

$$|a\rangle = \mathbf{\Lambda}_0 |i\rangle, \quad |a_1\rangle = \mathbf{\Lambda}_1 |i\rangle, \quad (\text{C1})$$

where $\mathbf{\Lambda}_0$ and $\mathbf{\Lambda}_1$ are corresponding transformation matrixes,

$$\mathbf{\Lambda}_0 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -i & 0 \\ 1 & i & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix}, \quad (\text{C2})$$

$$\mathbf{\Lambda}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} \cos \theta_1 \cos \varphi_1 + i \sin \varphi_1 & \cos \theta_1 \sin \varphi_1 - i \cos \varphi_1 & -\sin \theta_1 \\ \cos \theta_1 \cos \varphi_1 - i \sin \varphi_1 & \cos \theta_1 \sin \varphi_1 + i \cos \varphi_1 & -\sin \theta_1 \\ \sqrt{2} \sin \theta_1 \cos \varphi_1 & \sqrt{2} \sin \theta_1 \sin \varphi_1 & \sqrt{2} \cos \theta_1 \end{bmatrix}. \quad (\text{C3})$$

Using (C1), we can write

$$\begin{aligned} \langle a | a_1 \rangle &= \left(\sum_j [\Lambda_0]_{a,j} |j\rangle \right)^\dagger \sum_i [\Lambda_1]_{a_1,i} |i\rangle = \sum_{ij} \langle j | [\Lambda_0^\dagger]_{a,j} [\Lambda_1]_{a_1,i} |i\rangle \\ &= \sum_{ij} [\Lambda_0^\dagger]_{a,j} [\Lambda_1]_{a_1,i} \langle j | i \rangle = \sum_{ij} [\Lambda_0^\dagger]_{a,j} [\Lambda_1]_{a_1,i} \delta_{ij} = [\Lambda_1 \Lambda_0^\dagger]_{a_1,a} = [\Lambda_0^* \Lambda_1^T]_{a,a_1}. \end{aligned} \quad (\text{C4})$$

Substituting (C2) and (C3) into (C4), we obtain

$$\langle a | a_1 \rangle = \frac{1}{2} \begin{bmatrix} (\cos \theta_1 + 1)e^{i\varphi_1} & (\cos \theta_1 - 1)e^{i\varphi_1} & \sqrt{2} \sin \theta_1 e^{i\varphi_1} \\ (\cos \theta_1 - 1)e^{-i\varphi_1} & (\cos \theta_1 + 1)e^{-i\varphi_1} & \sqrt{2} \sin \theta_1 e^{-i\varphi_1} \\ -\sqrt{2} \sin \theta_1 & -\sqrt{2} \sin \theta_1 & 2 \cos \theta_1 \end{bmatrix}_{a,a_1}. \quad (\text{C5})$$

Using a similar procedure, we can obtain the transformation matrix $\langle a_2 | a \rangle$. However, since $|a_1\rangle$ and $|a_2\rangle$ are defined in similar fashion, the matrix $\langle a_2 | a \rangle$ can be obtained by taking the Hermitian adjoint of $\langle a | a_1 \rangle$, and replacing 1 with 2,

$$\langle a_2 | a \rangle = \frac{1}{2} \begin{bmatrix} (\cos \theta_2 + 1)e^{-i\varphi_2} & (\cos \theta_2 - 1)e^{i\varphi_2} & -\sqrt{2} \sin \theta_2 \\ (\cos \theta_2 - 1)e^{-i\varphi_2} & (\cos \theta_2 + 1)e^{i\varphi_2} & -\sqrt{2} \sin \theta_2 \\ \sqrt{2} \sin \theta_2 e^{-i\varphi_2} & \sqrt{2} \sin \theta_2 e^{i\varphi_2} & 2 \cos \theta_2 \end{bmatrix}_{a_2,a}. \quad (\text{C6})$$

-
- [1] J. Lee, M. Tymchenko, C. Argyropoulos, P.-Y. Chen, F. Lu, F. Demmerle, G. Boehm, M.-C. Amann, A. Alù, and M. A. Belkin, *Nature* **511**, 65 (2014).
- [2] S. Campione, A. Benz, M. B. Sinclair, F. Capolino, and I. Brener, *Appl. Phys. Lett.* **104**, 131104 (2014).
- [3] N. Segal, S. Keren-Zur, N. Hendler, and T. Ellenbogen, *Nat. Photonics* **9**, 180 (2015).
- [4] O. Wolf, S. Campione, A. Benz, A. P. Ravikumar, S. Liu, T. S. Luk, E. A. Kadlec, E. A. Shaner, J. F. Klem, M. B. Sinclair, and I. Brener, *Nat. Commun.* **6**, 7667 (2015).
- [5] A. Gabbay, J. Reno, J. R. Wendt, A. Gin, M. C. Wanke, M. B. Sinclair, E. Shaner, and I. Brener, *Appl. Phys. Lett.* **98**, 203103 (2011).
- [6] A. Benz, S. Campione, M. W. Moseley, J. J. Wierer, A. A. Allerman, J. R. Wendt, and I. Brener, *ACS Photonics* **1**, 906 (2014).
- [7] J. Lee, N. Nookala, J. S. Gomez-Diaz, M. Tymchenko, F. Demmerle, G. Boehm, M.-C. Amann, A. Alù, and M. A. Belkin, *Adv. Opt. Mater.* **4**, 664 (2016).
- [8] M. Kauranen, *Science* **342**, 1182 (2013).
- [9] M. Tymchenko, J. S. Gomez-Diaz, J. Lee, N. Nookala, M. A. Belkin, and A. Alù, *Phys. Rev. Lett.* **115**, 207403 (2015).
- [10] S. Keren-Zur, O. Avayu, L. Michaeli, and T. Ellenbogen, *ACS Photonics* **3**, 117 (2016).
- [11] E. Almeida, G. Shalem, and Y. Prior, *Nat. Commun.* **7**, 10367 (2016).
- [12] N. Nookala, J. Lee, M. Tymchenko, J. Sebastian Gomez-Diaz, F. Demmerle, G. Boehm, K. Lai, G. Shvets, M.-C. Amann, A. Alu, and M. Belkin, *Optica* **3**, 283 (2016).
- [13] Z. Bomzon, G. Biener, V. Kleiner, and E. Hasman, *Opt. Lett.* **27**, 1141 (2002).
- [14] L. Marrucci, C. Manzo, and D. Paparo, *Appl. Phys. Lett.* **88**, 221102 (2006).
- [15] D. Lin, P. Fan, E. Hasman, and M. L. Brongersma, *Science* **345**, 298 (2014).
- [16] M. Kang, T. Feng, H.-T. Wang, and J. Li, *Opt. Express* **20**, 15882 (2012).
- [17] S.-C. Jiang, X. Xiong, Y.-S. Hu, S.-W. Jiang, Y.-H. Hu, D.-H. Xu, R.-W. Peng, and M. Wang, *Phys. Rev. B* **91**, 125421 (2015).
- [18] X. Chen, L. Huang, H. Mühlenbernd, G. Li, B. Bai, Q. Tan, G. Jin, C.-W. Qiu, S. Zhang, and T. Zentgraf, *Nat. Commun.* **3**, 1198 (2012).
- [19] X. Ding, F. Monticone, K. Zhang, L. Zhang, D. Gao, S. N. Burokur, A. de Lustrac, Q. Wu, C.-W. Qiu, and A. Alù, *Adv. Mater.* **27**, 1195 (2015).
- [20] Z. Bomzon, V. Kleiner, and E. Hasman, *Opt. Lett.* **26**, 1424 (2001).
- [21] O. Avayu, O. Eisenbach, R. Ditcovski, and T. Ellenbogen, *Opt. Lett.* **39**, 3892 (2014).
- [22] S. Photonics, O. Society, A. Ocis, and O. Society, *Opt. Lett.* **40**, 3193 (2015).
- [23] Z. Li, K. Yao, F. Xia, S. Shen, J. Tian, and Y. Liu, *Sci. Rep.* **5**, 12423 (2015).
- [24] G. Biener, A. Niv, V. Kleiner, and E. Hasman, *Opt. Lett.* **27**, 1875 (2002).
- [25] P. Genevet, N. Yu, F. Aieta, J. Lin, M. A. Kats, R. Blanchard, M. O. Scully, Z. Gaburro, and F. Capasso, *Appl. Phys. Lett.* **100**, 013101 (2012).
- [26] G. Li, M. Kang, S. Chen, S. Zhang, E. Y. B. Pun, K. W. Cheah, and J. Li, *Nano Lett.* **13**, 4148 (2013).
- [27] L. Huang, X. Chen, H. Mühlenbernd, G. Li, B. Bai, Q. Tan, G. Jin, T. Zentgraf, and S. Zhang, *Nano Lett.* **12**, 5750 (2012).
- [28] C. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. (Wiley, Hoboken, New Jersey, 2005).
- [29] COMSOL, <https://www.comsol.com> (2015).
- [30] J. S. Gomez-Diaz, M. Tymchenko, J. Lee, M. A. Belkin, and A. Alù, *Phys. Rev. B* **92**, 125429 (2015).

- [31] L. Novotny, *Principles of Nano-Optics* (Cambridge University Press, New York, 2006).
- [32] N. Shitrit, S. Maayani, D. Veksler, V. Kleiner, and E. Hasman, *Opt. Lett.* **38**, 4358 (2013).
- [33] E. I. Rashba, *Sov. Phys. Solid State* **2**, 1224 (1960).
- [34] K. Ishizaka, M. S. Bahrmy, H. Murakawa, M. Sakano, T. Shimojima, T. Sonobe, K. Koizumi, S. Shin, H. Miyahara, A. Kimura, K. Miyamoto, T. Okuda, H. Namatame, M. Taniguchi, R. Arita, N. Nagaosa, K. Kobayashi, Y. Murakami, R. Kumai, Y. Kaneko, Y. Onose, and Y. Tokura, *Nat. Mater.* **10**, 521 (2011).
- [35] E. Rosencher, P. Bois, J. Nagle, and S. Delattre, *Electron. Lett.* **25**, 1063 (1989).
- [36] M. M. Fejer, S. J. B. Yoo, R. L. Byer, A. Harwit, and J. S. Harris, Jr., *Phys. Rev. Lett.* **62**, 1041 (1989).
- [37] E. Rosencher, A. Fiore, B. Vinter, V. Berger, Ph. Bois, and J. Nagle, *Science* **271**, 168 (1996).
- [38] F. Capasso, C. Sirtori, and A. Y. Cho, *IEEE J. Quantum Electron.* **30**, 1313 (1994).
- [39] C. Gmachl, A. Belyanin, D. L. Sivco, M. L. Peabody, N. Owschimikow, A. M. Sergent, F. Capasso, and A. Y. Cho, *IEEE J. Quantum Electron.* **39**, 1345 (2003).
- [40] M. A. Belkin, F. Capasso, A. Belyanin, D. L. Sivco, A. Y. Cho, D. C. Oakley, C. J. Vineis, and G. W. Turner, *Nat. Photonics* **1**, 288 (2007).
- [41] K. Vijayraghavan, Y. Jiang, M. Jang, A. Jiang, K. Choutagunta, A. Vizbaras, F. Demmerle, G. Boehm, M. C. Amann, and M. A. Belkin, *Nat. Commun.* **4**, 2021 (2013).
- [42] D. L. Andrews and M. Babiker, *The Angular Momentum of Light* (Cambridge University Press, Cambridge, 2013).
- [43] N. V. Bloch, K. Shemer, A. Shapira, R. Shiloh, I. Juwiler, and A. Arie, *Phys. Rev. Lett.* **108**, 233902 (2012).
- [44] S. Pancharatnam, *Proc. Indian Acad. Sci.* **44**, 247 (1956).
- [45] M. V. Berry, *J. Mod. Opt.* **34**, 1401 (1987).