# Metasurface-Refractive Hybrid Lens Modeling with Vector Field Physical Optics

**Ko-Han Shih [1] and C. Kyle Renshaw [1,2,3,*]**

[1] College of Optics and Photonics, CREOL, University of Central Florida, Orlando, FL 32816, USA; ko-han.shih@ucf.edu
[2] Department of Physics, University of Central Florida, Orlando, FL 32816, USA
[3] Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL 32816, USA
* Correspondence: kyle.renshaw@ucf.edu

# 1. Ray transmission coefficients

## 1.1 Phase correction

The core of the polarization ray tracing in Section 2.3.2. lies in the calculation of transmission coefficients, as shown in Equation (43). For uncoated refractive surfaces, these coefficients can be determined using Fresnel's equations. For coated surfaces, however, a dedicated electromagnetic solver for stratified systems, such as the TMM, is required to obtain these coefficients. Importantly, a phase correction is required to integrate TMM in polarization raytracing. This section provides further details about this process.

Referring to Figure S1(a), a ray is traced through a surface from an ambient material (e.g., air) to a glass material. The surface is coated with a thin film stack with total thickness $d$ and is assumed to be grown into ambient material. Given a ray launched at point O with input ray electric field phasor $E_O$ and is incident at an angle $\theta_{in}$ with respect to surface normal, the objective is to obtain electric field phasor $E_B$ at point B after transmission through the glass.

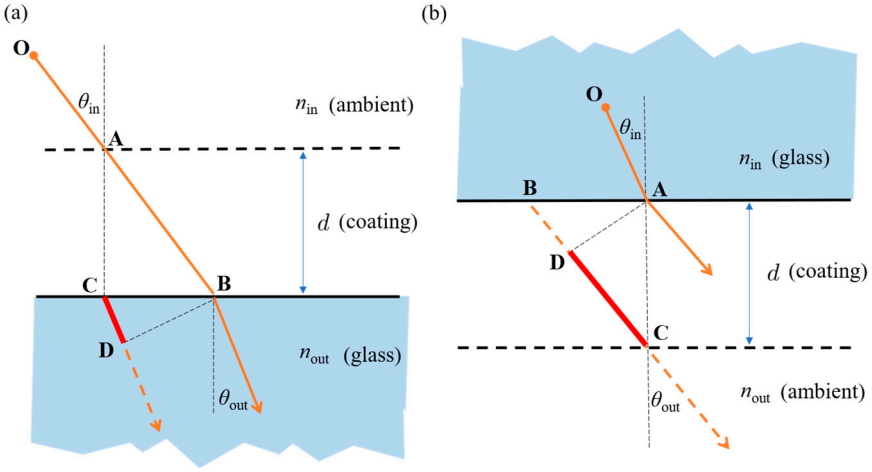



**Figure S1.** Illustrations for tracing a ray through a coated surface (a) from ambient material to glass and (b) from glass to ambient material.

In standard raytracing, the ray phase is accumulated along the ray path, and multiple reflections within a coating stack are not directly accounted for. Instead, a ray is first traced to a coated surface (e.g., point B) as if the coating were absent; then, transmission coefficients for the coating, with appropriate phase adjustments, are applied to update the ray's polarization. The TMM is commonly used to determine these transmission coefficients, relating electric fields between two points on either side of a coating along a line that is parallel to the surface normal (e.g., points A and C). Consequently, the output ray electric field phasor at point B is derived from the input ray phasor at point O by following these steps: (1) accumulate the optical path length $\overline{\mathrm{OB}}$, which is inherently provided by the ray tracing engine when the coating thickness is disregarded; (2) add the substrate optical path length $\overline{\mathrm{AB}}$ to obtain the ray phasor at point A; (3) multiply by the coating transmission coefficient $t$ to determine the ray phasor at point C; and (4) incorporate the optical path length $\overline{\mathrm{CD}}$ to calculate the ray phasor at point D, which is equivalent to the phasor at point B as the two points are on the same phase front. These procedures can be succinctly expressed as:

$$E_{\mathrm{B}} = E_{\mathrm{O}} \exp(ik_0 \mathrm{OPL}_{\overline{\mathrm{OB}}}) \exp(-ik_0 n_{\mathrm{in}} \overline{\mathrm{AB}})\, t \exp(ik_0 n_{\mathrm{out}} \overline{\mathrm{CD}})\,, \tag{S1}$$

where $n_{in}$ and $n_{out}$ are refractive indices of input and output regions, respectively, and $t$ is coating transmission coefficient, calculated from TMM, and relates electric fields between points A and C. From geometry and Snell's law we can find that $-k_0 n_{\mathrm{in}} \overline{\mathrm{AB}} + k_0 n_{\mathrm{out}} \overline{\mathrm{CD}} = -k_0 n_{\mathrm{in}} d \cos\theta_{\mathrm{in}}$. Thus, the formula for output ray electric field phasor is

$$E_{\mathrm{B}} = E_{\mathrm{O}} \exp(ik_0 \mathrm{OPL}_{\overline{\mathrm{OB}}})\, t \exp(-ik_0 n_{\mathrm{in}} d \cos\theta_{\mathrm{in}})\,, \tag{S2}$$

A similar equation can be obtained for rays traced from a glass material to ambient material. It is assumed the coating is growing into the ambient region. Refer to Figure S1(b), the electric field phasor at point A after considering coating transmission and phase adjustment is

$$E_A = E_O \exp(ik_0 \mathrm{OPL}_{\overline{OA}})\, t \exp(-ik_0 n_{out} \overline{CD})\,. \tag{S3}$$

From geometry we have $k_0 n_{out} \overline{CD} = k_0 n_{out} d \cos\theta_{out}$, so

$$E_A = E_O \exp(ik_0 \mathrm{OPL}_{\overline{OA}})\, t \exp(-ik_0 n_{out} d \cos\theta_{out})\,. \tag{S4}$$

Equations (S2) and (S4) suggest that, for a coated surface, depending on the ray orientation, the transmission coefficients $t_s$ and $t_p$ in Equation (40) should be modified as

$$\tilde{t}_{s(p)} = \begin{cases} t_{s(p)} \exp(-ik_0 n_{in} d \cos\theta_{in}) & \text{for incident ray from ambient material} \\ t_{s(p)} \exp(-ik_0 n_{out} d \cos\theta_{out}) & \text{else} \end{cases}. \tag{S5}$$

where $t_{s(p)}$ is s-(p-)polarized transmission coefficient calculated from TMM.

### 1.2 Ideal anti-reflection coating

In this work, an anti-reflection coating on a refractive surface is considered *ideal* if it results in no transmission loss when a polarized ray passes through the coated surface. This hypothetical coating has a transmission coefficient that depends on the angle of incidence:

$$t_s = t_p = \sqrt{\frac{n_{in} \cos\theta_{in}}{n_{out} \cos\theta_{out}}}, \tag{S6}$$

where $n_{in(out)}$ is refractive index of the input(output) region, and $\theta_{in(out)}$ is incident(refracted) angle. Equation (S6) can be substituted in Equation (43) in the main text for polarization raytracing through a refractive surface with ideal coating.

## 2. RFM versus GDM

In Section 3.1, the GDM was selected as the primary propagator for simulating field propagation through the refractive optics following the MS. Given the relatively short optical path length within the refractive elements, path diffraction effects are negligible in this context, allowing us to consider diffraction only for propagation beyond the refractive optics to the image plane. Additionally, the wavefront remains relatively simple, despite some aberrations, as it passes through the system. This simplicity allows for an alternative approach: using the RFM, as outlined in Section 2.4, to propagate the field through refractive optics. In Figure S2, we compare the maximum field angle PSF y-slices on the image plane of the hybrid lens in Figure 9(a) with and without the MS, using either GDM or RFM for propagation through the refractive elements. The agreement between the results validates the use of RFM when a simple wavefront is considered.

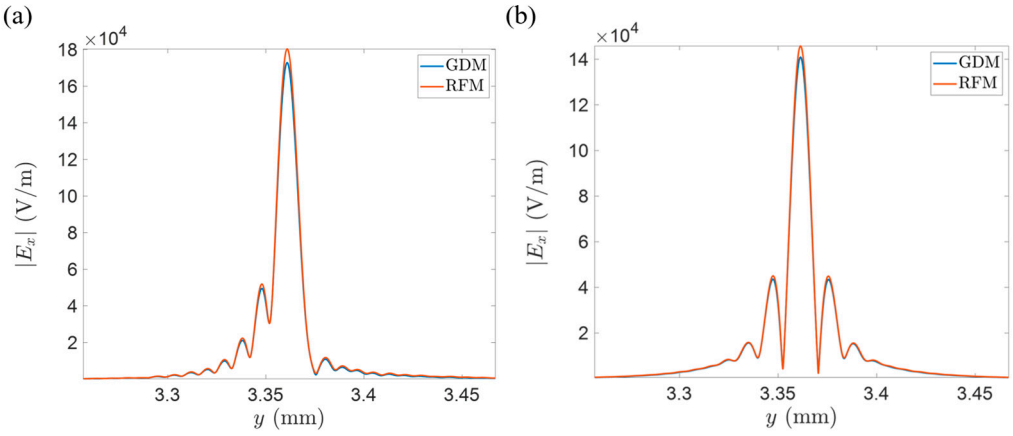



**Figure S2.** Comparison of maximum field angle PSF y-slices on the image plane of the imaging lens in Figure 9 (a) with and (b) without the MS. For each configuration, either GDM or RFM is used to propagate field through refractive optics.

## 3. Raytracing through a curved polarization-insensitive metasurface

This section details formulism for polarization raytracing through a curved polarization-insensitive MS, with its implementation demonstrated in Section 3.2. Referring to Figure S3(a), a MS with radius distribution $R(x,y)$ is built on a curved substrate with sag function $z(x,y)$. The MS imparts a phase gradient $\nabla\Phi(x,y;\lambda_0)$ on the xy-plane, where $\lambda_0$ is ray wavelength. This phase gradient can be estimated by mapping $R(x,y)$ to $\Phi(x,y;\lambda_0)$ using the meta-atom database, followed by numerical differentiation. A ray with direction $\hat{\mathbf{k}}_{\text{in}}$ and polarization $\mathbf{E}_{\text{in}}$ is incident on the curved MS, and the goal is to compute output ray direction $\hat{\mathbf{k}}_{\text{out}}$ and polarization $\mathbf{E}_{\text{out}}$.

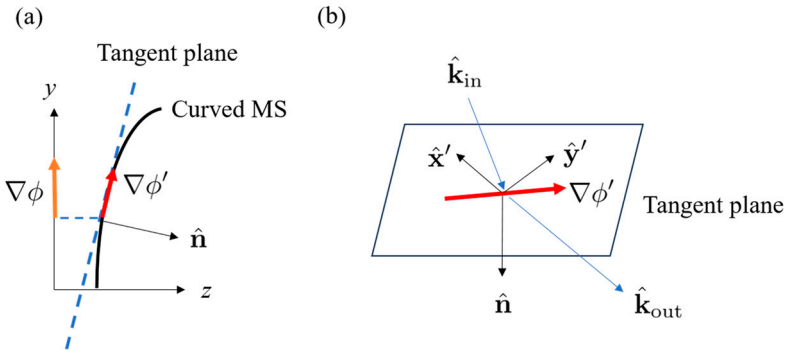



**Figure S3.** (a) Illustration for a curved MS. (b) Raytracing through a patch of a curved MS.

To begin, the incident ray is traced to the curved surface with intercept $(x_c, y_c, z_c)$, where a tangent plane with a surface normal $\hat{\mathbf{n}}$ can be established. Output ray direction is calculated as

$$\hat{\mathbf{k}}_{\text{out}} = \mu\left[\hat{\mathbf{k}}_{\text{in}} - \left(\hat{\mathbf{k}}_{\text{in}} \cdot \hat{\mathbf{n}}\right)\hat{\mathbf{n}}\right] + \frac{\lambda_0}{2\pi n_{\text{out}}}\nabla\Phi' + \sqrt{1 - \left|\mu\left[\hat{\mathbf{k}}_{\text{in}} - \left(\hat{\mathbf{k}}_{\text{in}} \cdot \hat{\mathbf{n}}\right)\hat{\mathbf{n}}\right] + \frac{\lambda_0}{2\pi n_{\text{out}}}\nabla\Phi'\right|^2}\,\hat{\mathbf{n}}, \tag{S7}$$

where $\mu = n_{\text{in}}/n_{\text{out}}$ with $n_{\text{in}}$ and $n_{\text{out}}$ are refractive index of the input and output region, respectively, and

$$\nabla\Phi' = \nabla\Phi - (\nabla\Phi \cdot \hat{\mathbf{n}})\hat{\mathbf{n}} \tag{S8}$$

is projection of phase gradient on the xy-plane to the tangent plane.

Next, we determine $\mathbf{E}_{\text{out}}$ using the meta-atom database. Since the database is typically constructed under the assumption that the MS surface normal is aligned with the z-direction, applying it to a local MS patch with an arbitrary normal requires a transformation into a local coordinate system:

$$\hat{\mathbf{x}}' = \frac{\hat{\mathbf{x}} + \frac{\partial z}{\partial x}\hat{\mathbf{z}}}{\sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2}}, \tag{S9a}$$

$$\hat{\mathbf{y}}' = \hat{\mathbf{n}} \times \hat{\mathbf{x}}', \tag{S9b}$$

with $\hat{\mathbf{x}}'$ and $\hat{\mathbf{y}}'$ are on the tangent plane. Note that $(\hat{\mathbf{x}}', \hat{\mathbf{y}}', \hat{\mathbf{n}})$ is a convenient basis to estmiate local zenith and azimuthal angles of the incident ray with respect to the tangent plane:

$$\theta = \cos^{-1}\left(\hat{\mathbf{k}}_{\text{in}} \cdot \hat{\mathbf{n}}\right), \tag{S10a}$$

$$\phi = \tan^{-1}\left(\frac{\hat{\mathbf{k}}_{\text{in}} \cdot \hat{\mathbf{y}}'}{\hat{\mathbf{k}}_{\text{in}} \cdot \hat{\mathbf{x}}'}\right). \tag{S10b}$$

Expressing the input polarization in the local coordinate system $\mathbf{E}_{\text{in}} = E_{x',\text{in}}\hat{\mathbf{x}}' + E_{y',\text{in}}\hat{\mathbf{y}}' + E_{n,\text{in}}\hat{\mathbf{n}}$, the transverse components for output ray polarization are calculated using the meta-atom database as

$$\begin{bmatrix} E_{x',\text{out}} \\ E_{y',\text{out}} \end{bmatrix} = \begin{bmatrix} t_{xx} & t_{xy} \\ t_{yx} & t_{yy} \end{bmatrix} \begin{bmatrix} E_{x',\text{in}} \\ E_{y',\text{in}} \end{bmatrix}, \tag{S11}$$

and the associated normal component is determined from the orthogonality condition

$$E_{n,\text{out}} = \frac{-(\hat{\mathbf{k}}_{\text{out}} \cdot \hat{\mathbf{x}}')E_{x',\text{out}} - (\hat{\mathbf{k}}_{\text{out}} \cdot \hat{\mathbf{y}}')E_{y',\text{out}}}{\hat{\mathbf{k}}_{\text{out}} \cdot \hat{\mathbf{n}}}. \tag{S12}$$

Finally, with output polarization in the local coordinate system $\mathbf{E}_{\text{out}} = E_{x',\text{out}}\hat{\mathbf{x}}' + E_{y',\text{out}}\hat{\mathbf{y}}' + E_{n,\text{out}}\hat{\mathbf{n}}$, a conversion back to the global coordinate system is applied.

Meanwhile, the polarization raytracing matrix across a curved MS is estimated as

$$\mathbf{P} = \mathbf{Y}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T, \tag{S13}$$

where $\mathbf{X} = [\mathbf{E}_{\text{in}} \quad \hat{\mathbf{k}}_{\mathbf{in}}]$ and $\mathbf{Y} = [\mathbf{E}_{\text{out}} \quad \hat{\mathbf{k}}_{\text{out}}]$ are 3-by-2 matrices with all elements expressed in the global coordinate system.

# 4. Hybrid lens for cylindrical polarization focusing

## 4.1 Polarization effect from an ideal thin lens

Consider a thin lens with focal length $f$ illuminated by a beam $\mathbf{E}_i$ that is transversely polarized with respect to the optical axis. At any given point in front of the thin lens, the incident field can be decomposed into two orthogonal components: a radial component and a tangential component. The unit vectors for these components are:

$$\hat{\mathbf{p}} = \cos\phi\,\hat{\mathbf{x}} + \sin\phi\,\hat{\mathbf{y}}, \tag{S14a}$$

$$\hat{\mathbf{s}} = -\sin\phi\,\hat{\mathbf{x}} + \cos\phi\,\hat{\mathbf{y}}, \tag{S14b}$$

where $\phi$ is the polar angle for a polar coordinate system center at the optical axis. Upon refraction by the thin lens, the tangential unit vector remains unchanged. However, the radial component is refracted to

$$\hat{\mathbf{r}} = \cos\theta\cos\phi\,\hat{\mathbf{x}} + \cos\theta\sin\phi\,\hat{\mathbf{y}} + \sin\theta\hat{\mathbf{z}}, \tag{S15}$$

which is the radial unit vector for a spherical coordinate system center at the focal point of the thin lens. In Reference [2], the transmitted beam is defined on a spherical reference surface centered at the focal point as

$$\mathbf{E}_t = (\mathbf{E}_i \cdot \hat{\mathbf{p}})\hat{\mathbf{r}} + (\mathbf{E}_i \cdot \hat{\mathbf{s}})\hat{\mathbf{s}}, \tag{S16}$$

and is followed by a Debye diffraction integral applied over the spherical reference surface to calculate field $\mathbf{E}_o$ near the focus. For more details, readers may consult the original literature. In this work, we streamline the calculation by directly performing a diffraction integral on a planar surface. We first add the phase of an ideal lens, resulting in an expression similar to that of the ASM:

$$\mathbf{E}_o = \mathcal{F}^{-1}\left\{\mathcal{F}\left\{\mathbf{E}_t \exp\left(-ik_0\sqrt{f^2 + x^2 + y^2}\right)\right\}\exp(ik_z f)\right\}. \tag{S17}$$

Equation (S17) is then used in the thin lens model in Section 3.3.

## 4.2 Azimuthally polarized beam focusing

In Section 3.3, the input field prior to the MS is x-polarized. The MS converts this field into a radially polarized beam, which is then focused by the doublet lens. Conversely, when the input beam is y-polarized, the MS generates an azimuthally polarized beam that is similarly focused by the same doublet lens. In this case, the focal spot is expected to be dominantly polarized in the azimuthal direction. The same ray-wave hybrid propagation framework discussed in the main text is used to simulate this behavior.

Figure S4 presents a comparison of the cylindrical components of the PSFs, obtained using both the ray-wave hybrid propagation through the hybrid lens and the ASM applied to its thin-lens model. The strong similarity between the two results further validates the accuracy of the GDM. As expected, the PSF exhibits a donut-shaped profile dominated by the azimuthal polarization component. Notably, in contrast to the case of radially polarized beam focusing, the longitudinal (z-) component of the field remains weak when focusing an azimuthally polarized beam [1].

All the aforementioned simulations consider on-axis incidence only, as both the MS and the off-the-shelf doublet lens are designed for optimal on-axis performance. Nonetheless, the proposed method is general and can be extended to off-axis scenarios. To demonstrate this, we consider a y-polarized beam incident at a small angle of 2.5°. Figure S5(a) shows the azimuthal component of the resulting PSF, and

Figure S5(b) compares the cylindrical field components along the y-axis. For this analysis, the cylindrical coordinate system is centered at the centroid of the PSF to account for the shift introduced by the off-axis incidence. It is clear the dominating component is still in the azimuthal component. However, PSF quality degrades due to factors such as off-axis aberrations.

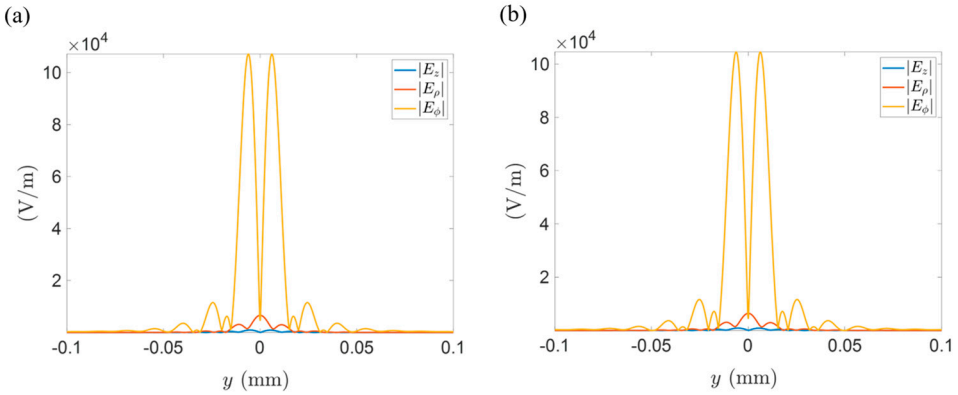



**Figure S4.** Magnitude of beam cylindrical component y-slices on the image plane for (a) the hybrid lens in Figure 13(a) and (b) its thin-lens model in Figure 14(a). The input field is on-axis and y-polarized.

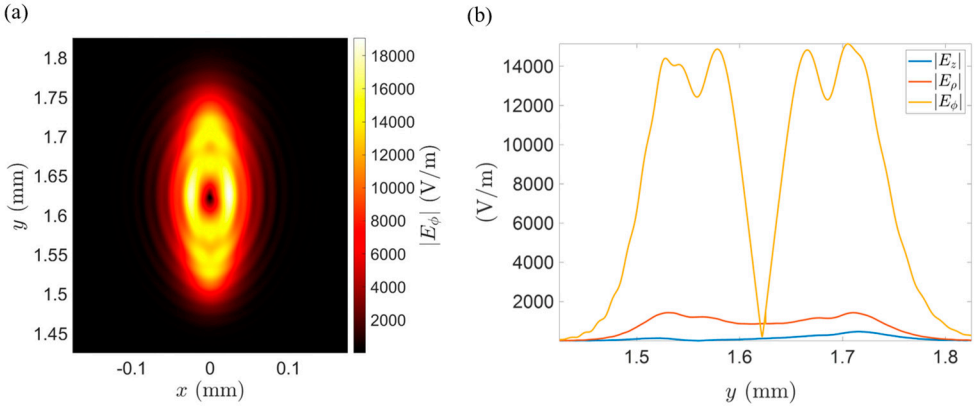



**Figure S5.** (a) Magnitude of PSF azimuthal component for the hybrid lens in Figure 13(a). (b) Magnitude of beam cylindrical component y-slices. The input field is off-axis and y-polarized.

# 5. Vector adjoint gradient method

This section provides the theoretical framework for the vector field AGM discussed in Section 3.4 and can be considered as a generalization for the scalar field AGM discussed in Reference [2]. Consider a MS-refractive hybrid system with multiple MSs. An input field is prepared on the input plane and sequentially propagated through the optical components to an output plane. The goal is to optimize the MS layouts such that the realized output field closely matches a specified target field on the output plane.

Using bra-ket notation for compact representation, the target field's x- and y-components are expressed as a composite ket:

$$\left|\mathbf{E}^{(\mathrm{tar})}\right\rangle = \begin{bmatrix} \left|E_x^{(\mathrm{tar})}\right\rangle \\ \left|E_y^{(\mathrm{tar})}\right\rangle \end{bmatrix}. \tag{S18}$$

where $\left|E_x^{(\mathrm{tar})}\right\rangle$ is effectively a vector of target field x-component at different positions on the output plane, and similar for $\left|E_y^{(\mathrm{tar})}\right\rangle$. Similarly, the output field achieved by the system is denoted as:

$$\left|\mathbf{E}^{(\mathrm{out})}\right\rangle = \begin{bmatrix} \left|E_x^{(\mathrm{out})}\right\rangle \\ \left|E_y^{(\mathrm{out})}\right\rangle \end{bmatrix}. \tag{S19}$$

To quantify the similarity between the target and output fields, the objective function $C$ is defined as:

$$C = \left|\left\langle \mathbf{E}^{(\mathrm{tar})} \middle| \mathbf{E}^{(\mathrm{out})} \right\rangle\right|^2. \tag{S20}$$

which is proportional to the squared inner product between the target and output fields.

For gradient decent optimization, consider the $m$-th MS in the system, where each meta-atom has a tunable parameter $p$. The layout of these parameters across the MS is denoted as $\mathbf{p}^{(m)}$. To optimize the MS layout, it is necessary to compute the gradient of the objective function with respect to the parameters, $\partial C/\partial \mathbf{p}^{(m)}$, at every position on the MS.

As will be demonstrated, the AGM offers an efficient approach to calculate $\partial C/\partial \mathbf{p}^{(m)}$ for all positions simultaneously. This efficiency is achieved by combining the forward propagation of the input field through the optical system with the backward propagation of an adjoint field. The latter's expression is derived in the following discussion.

Consider a specific position s′ on the $m$-th MS, the gradient is expanded as:

$$\frac{\partial C}{\partial p^{(m,s')}} = 2\mathrm{Re}\left\{ \left\langle \mathbf{E}^{(\mathrm{out})} \middle| \mathbf{E}^{(\mathrm{tar})} \right\rangle \left\langle \mathbf{E}^{(\mathrm{tar})} \left| \frac{\partial}{\partial p^{(m,s')}} \right| \mathbf{E}^{(\mathrm{out})} \right\rangle \right\}, \tag{S21}$$

where the formula $\nabla|F|^2 = 2\mathrm{Re}\{F^*\nabla F\}$ is applied, with $F$ being a complex-valued quantity. In Equation (S21), the first term inside the brace, $\left\langle \mathbf{E}^{(\mathrm{out})} \middle| \mathbf{E}^{(\mathrm{tar})} \right\rangle$, can be determined through the forward propagation of the input field to the output plane. Therefore, the remaining task is to compute the second term $\left\langle \mathbf{E}^{(\mathrm{tar})} \left| \frac{\partial}{\partial p^{(m,s')}} \right| \mathbf{E}^{(\mathrm{out})} \right\rangle$, which involves the gradient of the output field with respect to the parameter $\partial \left|\mathbf{E}^{(\mathrm{out})}\right\rangle / \partial p^{(m,s')}$.

Since a small change in $p^{(m,s')}$ does not affect the fields before the $m$-th MS, the output field can be expressed as

$$\left|\mathbf{E}^{(\mathrm{out})}\right\rangle = \underbrace{\begin{bmatrix} \mathbf{p}_{xx}^{(m+)} & \mathbf{p}_{xy}^{(m+)} \\ \mathbf{p}_{yx}^{(m+)} & \mathbf{p}_{yy}^{(m+)} \end{bmatrix}}_{\mathbf{P}^{(m+)}} \underbrace{\begin{bmatrix} \mathbf{t}_{xx}^{(m+)} & \mathbf{t}_{xy}^{(m+)} \\ \mathbf{t}_{yx}^{(m+)} & \mathbf{t}_{yy}^{(m+)} \end{bmatrix}}_{\mathbf{T}^{(m)}} \underbrace{\begin{bmatrix} \left|E_x^{(m-)}\right\rangle \\ \left|E_y^{(m-)}\right\rangle \end{bmatrix}}_{\left|\mathbf{E}^{(m-)}\right\rangle}. \tag{S22}$$

Here, $\left|\mathbf{E}^{(m-)}\right\rangle$ represents field immediately in front of the $m$-th MS, $\mathbf{T}^{(m)}$ is the composite transmission matrix for the MS, and $\mathbf{P}^{(m+)}$ is the propagation matrix modeling the field progression through the subsequent optical elements to the output plane. In practice, $\mathbf{P}^{(m+)}$ is treated symbolically, as its explicit matrix elements are not computed directly. Instead, the propagation methods outlined in Section 2 of the main text are employed to numerically propagate the field through the different types of optical elements in the system. With Equation (S22), we can expand the second term in the brace of Equation (S21) as

$$\left\langle \mathbf{E}^{(\mathrm{tar})} \left| \frac{\partial}{\partial p^{(m,s')}} \right| \mathbf{E}^{(\mathrm{out})} \right\rangle = \left\langle \mathbf{E}^{(\mathrm{tar})} \left| \mathbf{P}^{(m+)} \frac{\partial}{\partial p^{(m,s')}} \mathbf{T}^{(m)} \right| \mathbf{E}^{(m-)} \right\rangle, \tag{S23}$$

which can be written in slightly different way to highlight the expression for adjoint field:

$$\left\langle \mathbf{E}^{(\mathrm{tar})} \left| \frac{\partial}{\partial p^{(m,s')}} \right| \mathbf{E}^{(\mathrm{out})} \right\rangle = \left[ \left(\mathbf{P}^{(m+)}\right)^\dagger \left|\mathbf{E}^{(\mathrm{tar})}\right\rangle \right]^\dagger \frac{\partial}{\partial p^{(m,s')}} \mathbf{T}^{(m)} \left|\mathbf{E}^{(m-)}\right\rangle. \tag{S24}$$

We define adjoint field immediately after the $m$-th MS as

$$\left|\mathbf{a}^{(m+)}\right\rangle = \begin{bmatrix} \left|a_x^{(m+)}\right\rangle \\ \left|a_y^{(m+)}\right\rangle \end{bmatrix} = \left(\mathbf{P}^{(m+)}\right)^\dagger \left|\mathbf{E}^{(\mathrm{tar})}\right\rangle, \tag{S25}$$

From the equation, we arrive at the critical conclusion that $\left|\mathbf{a}^{(m+)}\right\rangle$ is calculated simply by backward propagation of the target field to the plane immediately after the $m$-th MS.

Meanwhile, the gradient of the transmission matrix with respect to $p^{(m,s')}$ is expressed as:

$$\frac{\partial \mathbf{T}^{(m)}}{\partial p^{(m,s')}} = \frac{\partial}{\partial p^{(m,s')}} \begin{bmatrix} t_{xx}^{(m,s')} & t_{xy}^{(m,s')} \\ t_{yx}^{(m,s')} & t_{yy}^{(m,s')} \end{bmatrix} |s'\rangle\langle s'|, \tag{S26}$$

where $t_{ij}^{(m,s')}$ are the transmission coefficients of the $m$-th MS at the position $s'$, $|s'\rangle$ is position basis at $s'$. To derive Equation (S26), we assume that meta-atoms are independent of one another, meaning that a small variation at position $s'$ does not affect transmission coefficients at other positions. This independence simplifies the calculation of gradients, allowing localized adjustments to meta-atom parameters during optimization. Substituting Equations (S25) and (S26) into Equation (S24), we have

$$\left\langle \mathbf{E}^{(\mathrm{tar})} \left| \frac{\partial}{\partial p^{(m,s')}} \right| \mathbf{E}^{(\mathrm{out})} \right\rangle = \left\langle \mathbf{a}^{(m+,s')} \left| \frac{\partial}{\partial p^{(m,s')}} \begin{bmatrix} t_{xx}^{(m,s')} & t_{xy}^{(m,s')} \\ t_{yx}^{(m,s')} & t_{yy}^{(m,s')} \end{bmatrix} \right| \mathbf{E}^{(m-,s')} \right\rangle, \tag{S27}$$

with $|\mathbf{a}^{(m+,s')}\rangle$ and $|\mathbf{E}^{(m-,s')}\rangle$ being fields at the s′ position under consideration. Given Equation (S27), the objective function gradient for every position on a MS can be calculated from Equation (S21).

Importantly, all quantities in Equation (S27) are associated with the s′ position only. This means that calculation of objective function gradient for a parameter can be broadcasted to *all* positions on a MS, leading to

$$\frac{\partial C}{\partial \mathbf{p}^{(m)}} = 2\mathrm{Re}\left\{ \left\langle \mathbf{E}^{(\mathrm{out})} \middle| \mathbf{E}^{(\mathrm{tar})} \right\rangle \left\langle \mathbf{E}^{(\mathrm{tar})} \left| \frac{\partial}{\partial \mathbf{p}^{(m)}} \right| \mathbf{E}^{(\mathrm{out})} \right\rangle \right\}, \tag{S28}$$

with

$$\begin{aligned} \left\langle \mathbf{E}^{(\mathrm{tar})} \left| \frac{\partial}{\partial \mathbf{p}^{(m)}} \right| \mathbf{E}^{(\mathrm{out})} \right\rangle = & \left( \frac{\partial t_{xx}^{(m)}}{\partial p^{(m)}} E_x^{(m-)} + \frac{\partial t_{xy}^{(m)}}{\partial p^{(m)}} E_y^{(m-)} \right) \left(a_x^{(m+)}\right)^* \\ & + \left( \frac{\partial t_{yx}^{(m)}}{\partial p^{(m)}} E_x^{(m-)} + \frac{\partial t_{yy}^{(m)}}{\partial p^{(m)}} E_y^{(m-)} \right) \left(a_y^{(m+)}\right)^*. \end{aligned} \tag{S29}$$

For the righthand side of Equation (S29), it is understood that all the quantities are arrays on a surface and all operations (multiplications, summations, partial derivatives, etc.) are elementwise across arrays. By combining the forward propagation of the incident field and the backward propagation of the adjoint field, the objective function gradients for all parameters on all MSs can be efficiently computed. This process requires the partial derivatives of the transmission coefficients with respect to the meta-atom parameters, $\partial t_{ij}/\partial p$. These derivatives can be obtained, for example, by interpolation from a precomputed meta-atom polarization database.

For the design example presented in Section 3.4, the meta-atom is parameterized by three variables: principal transmission coefficients $\phi_{xx}^{(0)}$ and $\phi_{yy}^{(0)}$, and rotation angle $\theta_r$. The associated transmission matrix is expressed ass:

$$\begin{bmatrix} t_{xx} & t_{xy} \\ t_{yx} & t_{yy} \end{bmatrix} = \begin{bmatrix} \cos\theta_r & -\sin\theta_r \\ \sin\theta_r & \cos\theta_r \end{bmatrix} \begin{bmatrix} \exp\left(i\phi_{xx}^{(0)}\right) & 0 \\ 0 & \exp\left(i\phi_{yy}^{(0)}\right) \end{bmatrix} \begin{bmatrix} \cos\theta_r & \sin\theta_r \\ -\sin\theta_r & \cos\theta_r \end{bmatrix}. \tag{S30}$$

Analytical expressions for the partial derivatives of the transmission matrix elements with respect to the parameters are given as follows:

$$\begin{bmatrix} \frac{\partial t_{xx}}{\partial \theta_r} & \frac{\partial t_{xy}}{\partial \theta_r} \\ \frac{\partial t_{yx}}{\partial \theta_r} & \frac{\partial t_{yy}}{\partial \theta_r} \end{bmatrix} = \left[\exp\left(i\phi_{xx}^{(0)}\right) - \exp\left(i\phi_{yy}^{(0)}\right)\right] \begin{bmatrix} -\sin 2\theta_r & \cos 2\theta_r \\ \cos 2\theta_r & \sin 2\theta_r \end{bmatrix}, \tag{S31a}$$

$$\begin{bmatrix} \frac{\partial t_{xx}}{\partial \phi_{xx}^{(0)}} & \frac{\partial t_{xy}}{\partial \phi_{xx}^{(0)}} \\ \frac{\partial t_{yx}}{\partial \phi_{xx}^{(0)}} & \frac{\partial t_{yy}}{\partial \phi_{xx}^{(0)}} \end{bmatrix} = i \exp\left(i\phi_{xx}^{(0)}\right) \begin{bmatrix} \cos^2\theta_r & \frac{\sin 2\theta_r}{2} \\ \frac{\sin 2\theta_r}{2} & \sin^2\theta_r \end{bmatrix}, \tag{S31b}$$

$$\begin{bmatrix} \frac{\partial t_{xx}}{\partial \phi_{yy}^{(0)}} & \frac{\partial t_{xy}}{\partial \phi_{yy}^{(0)}} \\ \frac{\partial t_{yx}}{\partial \phi_{yy}^{(0)}} & \frac{\partial t_{yy}}{\partial \phi_{yy}^{(0)}} \end{bmatrix} = i \exp\left(i\phi_{yy}^{(0)}\right) \begin{bmatrix} \sin^2\theta_r & -\frac{\sin 2\theta_r}{2} \\ -\frac{\sin 2\theta_r}{2} & \cos^2\theta_r \end{bmatrix}. \tag{S31c}$$

These expressions, alongside Equations (S28) and (S29), form the basis for the adjoint gradient optimization of the MS parameters discussed in Section 3.4.

# 6. Full-Stokes polarization camera

## 6.1 Polarization-splitting MS design

This section provides the steps for designing the Full-Stokes polarization camera described in Section 3.5, using MS design strategies inspired by Reference [3]. As shown in Figure S6, the design begins with specifying four polarization states $\mathbf{J}_i$, $i = \{1, 2, 3, 4\}$, arranged as vertices of a regular tetrahedron on the Poincaré sphere. Each polarization state corresponds to a specific quadrant on the sensor super-cell, where its associated energy should be deflected.

A primary lens with a given F-number FN is used to focus the incident beam. Positioned near the focal plane of the primary lens, the MS serves dual functions: relaying the focused beam and separating its polarization components. The MS is characterized by two adjustable parameters: the pitch size of its super-cell $\Lambda$, and the back focal distance $f_{\text{MS}}$. It is assumed that the sensor super-pixel pitch matches the MS super-cell pitch. In Section 3.5, these parameters are set to $\Lambda = 40$ μm and $f_{\text{MS}} = 30$ μm.

The MS super-cell is defined by a Jones matrix $\mathbf{J}(x, y)$, which encodes the desired polarization-splitting functionality. To derive this matrix, the design considers the ideal Jones matrix that focuses a specific polarization state to the center of its corresponding quadrant. Using bra-ket notation, it is

$$\mathbf{J}_i(x, y) = \Phi_i(x, y)|\mathbf{J}_i^*\rangle\langle\mathbf{J}_i|, \tag{S32}$$

where $\mathbf{J}_i^*$ is the conjugate state for $\mathbf{J}_i$, and

$$\Phi_i(x, y) = \Phi_-(x, y)\Phi_{+i}(x, y), \tag{S33}$$

is the desired phase profile with

$$\Phi_-(x, y) = \exp\left(ik_0\sqrt{x^2 + y^2 + (\text{FN} \cdot \Lambda)^2}\right), \tag{S34}$$

being the phase profile with a negative optical power to collimate a converged beam after the primary lens, and

$$\Phi_{+i}(x, y) = \exp\left(-ik_0\sqrt{(x - x_i)^2 + (y - y_i)^2 + f_{\text{MS}}^2}\right), \tag{S35}$$

being the phase profile with a positive optical power to focus a collimated beam to the designated quadrant center $(x_i, y_i)$ on the sensor super-cell. Note that Equation (S34) assumes a converged beam after the primary lens has its chief ray parallel to the optical axis, meaning the primary lens should be designed to have small chief ray angles across the FOV in the image space (i.e., telecentric in image space).

The ideal Jones matrix for the MS super-cell is

$$\mathbf{J}_t(x, y) = \frac{1}{\sqrt{2}} \sum_{i=1}^{4} \mathbf{J}_i(x, y), \tag{S36}$$

Note the matrix in Equation (S36) is symmetric but *not* unitary and thus cannot be realized by meta-atoms described by Equation (59) in the main text, which is symmetric *and* unitary. Following Reference [4], applying polar decomposition to Equation (S36) such that

$$\mathbf{J}_t(x, y) = \mathbf{P}(x, y)\mathbf{U}(x, y), \tag{S37}$$

where $\mathbf{P}$ is positive semi-definite and $\mathbf{U}$ is unitary, and the approximation $\mathbf{J}(x,y) := \mathbf{U}(x,y)$ is made. Finally, to determine the associated MS parameters $\phi_{xx}^{(0)}$, $\phi_{yy}^{(0)}$, and $\theta_r$ that can realize $\mathbf{J}(x,y)$, eigen-decomposition is applied such that

$$\mathbf{J}(x,y) = \mathbf{O}(x,y)\mathbf{\Phi}(x,y)\mathbf{O}^{-1}(x,y), \tag{S38}$$

with

$$\mathbf{\Phi}(x,y) = \begin{bmatrix} \exp\left(i\phi_{xx}^{(0)}(x,y)\right) & 0 \\ 0 & \exp\left(i\phi_{yy}^{(0)}(x,y)\right) \end{bmatrix}, \tag{S39a}$$

$$\mathbf{O}(x,y) = \begin{bmatrix} \cos\theta_r(x,y) & -\sin\theta_r(x,y) \\ \sin\theta_r(x,y) & \cos\theta_r(x,y) \end{bmatrix}. \tag{S39b}$$

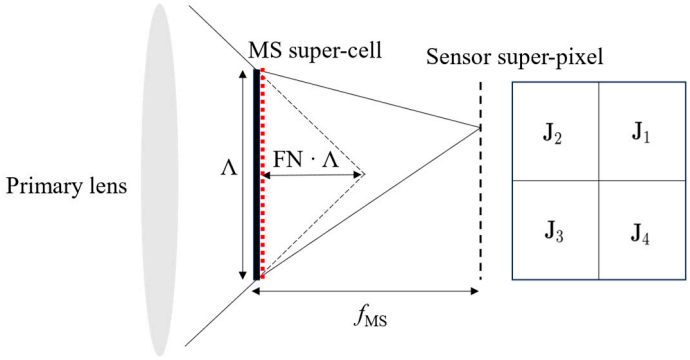



**Figure S6**. Illustration for polarization-splitting MS design.

## 6.2 Device matrix

When an input beam with an unknown polarization reaches the image plane of the polarization camera, the intensities integrated across the four quadrants yield a power vector, $\mathbf{P}$. A device matrix, $\mathbf{B}$, is then applied to estimate the incident Stokes vector as $\mathbf{S} = \mathbf{BP}$. In the main text, the device matrix was assumed to be calibrated for on-axis incidence. This calibration is performed by preparing four on-axis beams, each corresponding to one of the designated polarization states $\mathbf{J}_i$ ($i$ = {1, 2, 3, 4}). For each beam, the known Stokes vector $\mathbf{S}_i$ is paired with the measured power vector $\mathbf{P}_i$ on the image plane. The device matrix is then computed as:

$$\mathbf{B} = [\mathbf{S}_1 \quad \mathbf{S}_2 \quad \mathbf{S}_3 \quad \mathbf{S}_4][\mathbf{P}_1 \quad \mathbf{P}_2 \quad \mathbf{P}_3 \quad \mathbf{P}_4]^{-1}. \tag{S40}$$

This device matrix, calibrated for on-axis incidence (Figure S7(a)), is applied to estimate the polarization for off-axis beams. However, as illustrated in Figure 19 in the main text, the accuracy of polarization estimation degrades for off-axis fields.

To address this problem, angular-variant device matrices can be generated for different regions of the FOV, thereby improving accuracy for off-axis beams. As an example, a device matrix calibrated at the maximum field angle of 30° was created using incident beams at this angle. This calibration produces a modified device matrix (Figure S7(b)), which significantly improves the accuracy of polarization estimation for beams at the maximum field angle, as demonstrated in Figure S7(c).

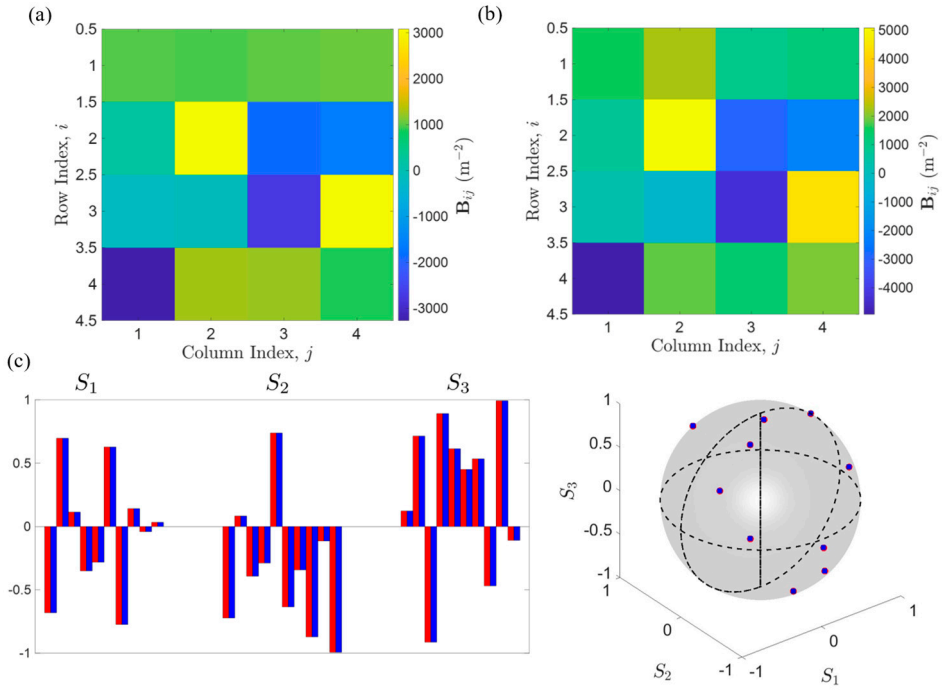



**Figure S7.** Device matrix calibrated for (a) on-axis incidence and (b) for incident angle of 30°. (c) Comparisons of Stokes vectors for the detected states (blue) and the incident states (red) for an incident angle of 30°.

## References


[1] Zhan, Q.; Leger, J. R. Focus Shaping Using Cylindrical Vector Beams. *Opt. Express* **2002**, *10*, 324–331. https://doi.org/10.1364/OE.10.000324.

[2] Shih, K.-H.; Renshaw, C. K. Hybrid Meta/Refractive Lens Design with an Inverse Design Using Physical Optics. *Appl. Opt.* **2024**, *63*, 4032–4043. https://doi.org/10.1364/AO.516890.

[3] Soma, G.; Komatsu, K.; Ren, C.; Nakano, Y.; Tanemura, T. Metasurface-Enabled Non-Orthogonal Four-Output Polarization Splitter for Non-Redundant Full-Stokes Imaging. *Opt. Express* **2024**, *32*, 34207–34222. https://doi.org/10.1364/OE.529389.