Transcending shift-invariance in the paraxial regime via end-to-end inverse design of freeform nanophotonics

William F. Li,[1,*] **Gaurav Arya,**[2] **Charles Roques-Carmes,**[1] **Zin Lin,**[3] **Steven G. Johnson,**[2] **Marin Soljačić**[1,4]

[1]*Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139*
[2]*Department of Mathematics, Massachusetts Institute of Technology, Cambridge, MA 02139*
[3]*Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA 24060*
[4]*Department of Physics, Massachusetts Institute of Technology, Cambridge, MA 02139*
[*]*wfli@mit.edu*

Abstract: Traditional optical elements and conventional metasurfaces obey shift-invariance in the paraxial regime. For imaging systems obeying paraxial shift-invariance, a small shift in input angle causes a corresponding shift in the sensor image. Shift-invariance has deep implications for the design and functionality of optical devices, such as the necessity of free space between components (as in compound objectives made of several curved surfaces). We present a method for nanophotonic inverse design of compact imaging systems whose resolution is not constrained by paraxial shift-invariance. Our method is end-to-end, in that it integrates density-based full-Maxwell topology optimization with a fully iterative elastic-net reconstruction algorithm. By the design of nanophotonic structures that scatter light in a non-shift-invariant manner, our optimized nanophotonic imaging system overcomes the limitations of paraxial shift-invariance, achieving accurate, noise-robust image reconstruction beyond shift-invariant resolution.

1. Introduction

The design of imaging systems that transcend paraxial shift-invariance is the next step in making compact, high-resolution imagers. Conventional optical elements, such as thin lenses [1], obey the property of paraxial shift-invariance, meaning the best angular resolution they can achieve is $\Delta\theta \sim \Delta x_D/d$, where d is the distance between the optical element and the sensor and Δx_D is the width of a detector pixel on the sensor (Fig 1a, left and right). Traditional metasurfaces relying on pre-computed paraxial phase libraries [2–7] are also constrained by the same limitation. Here, we present a method for the design of freeform nanophotonic optical elements that overcome such constraints on angular resolution. We demonstrate our method to design two-dimensional (2D) and three-dimensional (3D) freeform nanophotonic structures for angle-resolved spectrometry at angular resolutions beyond what is allowed by the paraxial limit. We design our freeform nanophotonic structures through topology optimization [8–10] in an end-to-end [11–16] pipeline (Fig. 1(d)), which directly minimizes the ultimate reconstruction error. In our approach, freeform nanophotonic geometries are parametrized by dielectric permittivity ε at every pixel in a design region, amounting to tens of thousands of optimization parameters. We show that our optimized structures outperform both a conventional thin lens (which obeys paraxial shift-invariance) and random nanophotonic structures (which are not beholden to the same limit).

Previously, the challenge of transcending paraxial shift-invariance has been addressed with super-cell metasurfaces [17] and nonlocal optics [2, 18–20], in the context of "space squeezing" – compression of free space by designed nanophotonic structures. For a given detector pixel size Δx_D in traditional optical elements, the paraxial shift-invariance limit dictates the distance d between the optical element and sensor required to capture images at a given spatial resolution $\Delta\theta$, where higher resolutions with $\Delta\theta < \Delta x_D/d$ cannot be captured by an imaging system
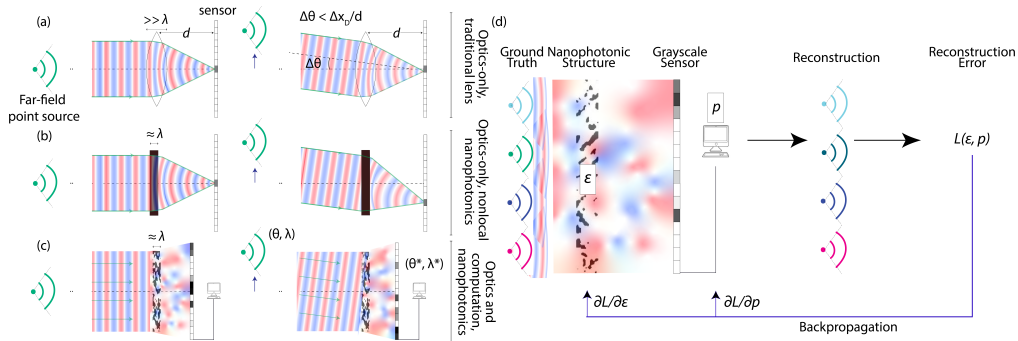
Fig. 1. **Transcending shift-invariance with end-to-end optimized freeform meta-surfaces.** (a) Left (resp. Right): Image formation with a conventional thin lens at normal (resp. oblique) incidence. The two angles are not resolved. (b) Left (resp. Right): Image formation with a designed nonlocal nanophotonic optical element at normal (resp. oblique) incidence. The two angles are resolved. (c) Left (resp. Right): Image formation in an end-to-end nanophotonic pipeline at normal (resp. oblique) incidence. The two angles are resolved with computational processing. (d) The end-to-end design pipeline presented in this work. From left to right: the ground truth emits polychromatic light scattered through a nanophotonic structure and measured on a grayscale sensor. The signal on the sensor is then fed to a reconstruction algorithm with hyperparameters p, which outputs a reconstruction of the original signal. During training, a reconstruction error $L(\varepsilon, p)$ is computed as a function of the nanophotonic structure permittivity profile ε and the reconstruction hyperparameters p. To design the structure ε, gradients $\partial L/\partial \varepsilon$ are backpropagated back to the nanophotonic structure. To tune the hyperparameters p, gradients $\partial L/\partial p$ are propagated back to the reconstruction algorithm.

with a conventional thin lens (Fig. 1(a)). Minimizing free space in imaging systems, or "space squeezing," is part of a broader effort to reduce the volume of imaging systems [21–23]. Prior work seeks to minimize the volume of both the optical element [3–6] and free-space [2, 18] separately, and typically involves two engineered structures: a local metasurface to replace the lens and a nonlocal structure to replace free space (e.g. a multi-layer stack acting as a space squeezer). This two-structure system is necessary because traditional metasurfaces are characterized by local transfer functions while replacing free space requires a nonlocal (momentum-dependent) transfer function [2]. Nonlocal optimized metasurfaces can resolve angles within the paraxial regime, as shown in Fig. 1(b). Our method, shown in Fig. 1(c), offers a more compact alternative to space squeezing by designing a single thin engineered nanophotonic structure which, in conjunction with a computational-imaging algorithm, replaces both the lens and free space, rather than one for each. Unlike prior approaches that separate the two problems (lens and free space), our approach combines them into a single end-to-end image-reconstruction problem.

In addition to the nanophotonic structure design, a robust computational reconstruction component is essential to imaging beyond the paraxial limit. Prior work realized compact imagers in inverse-designed optics-only systems [24]. Such techniques rely on the optimization of many degrees of freedom (typically distributed over an entire optimization volume) to realize a pre-defined optical functionality. In contrast, our work leverages recent developments in end-to-end inverse-design: harnessing computational reconstruction to loosen constraints on the transformation imparted by the optimized optical elements. We achieve this by resolving images and patterns that would be unreadable to the human eye (and could not be pre-defined by the user as an optimization task). This allows for the design of thinner, less complex nanophotonic

structures that need only to produce an image interpretable by the reconstruction algorithm. In previous work in end-to-end optimization, nanophotonic structures have been paired with various image reconstruction algorithms, including neural networks, compressed sensing (or Lasso-regularized regression), Tikhonov-regularized regression, and elastic-net regression [25–29]. In our work, we use the elastic-net reconstruction algorithm, which combines the Lasso (l_1) regularization term and the Tikhonov (l_2) regularization term. Intuitively, the l_1 term encourages the regression to deliver a sparse solution, which is of particular interest to us in the detection of angle and frequency of incoming laser beams under the often reasonable assumption of there coming only a few beams at a time. We give the reconstruction algorithm flexibility to choose whether to emphasize the Lasso or Tikhonov regularization terms by optimizing the elastic-net hyperparameters [30]; we show that a sparse problem generally results in the optimization greatly emphasizing the Lasso term and de-emphasizing the Tikhonov term. Our reconstruction algorithm is paired with the topology-optimized structure, allowing for the automated discovery of both freeform designs and reconstruction hyperparameters.

2. End-to-end optimization pipeline

2.1. Image formation

We consider polychromatic and spatially-extended objects, therefore describing the ground-truth as a multi-dimensional tensor with at most 4 dimensions (3D space + 1D spectral). For convenience, we represent this tensor as a flattened single vector $\mathbf{u}$ where each component corresponds to a unique angle-frequency pair (θ, λ). We propagate the ground truth object through our nanophotonic structure and free space to generate a raw, noisy, grayscale image at the detector $\mathbf{v}$, where

$$\mathbf{v} = G\left(\varepsilon(\mathbf{r})\right)\mathbf{u} + \eta. \tag{1}$$

In the above expression, $\varepsilon(\mathbf{r})$ is the dielectric profile of the nanophotonic structure, $G\left(\varepsilon(\mathbf{r})\right)$ is the measurement matrix of the imaging system (a function of $\varepsilon(\mathbf{r})$), and η is the additive Gaussian noise, whose standard deviation is proportional to the average intensity on the detector. The degrees of freedom of the structure $\varepsilon(\mathbf{r})$ are free to take on arbitrary designs through the optimization, and, in particular, we do not assume that G should result in a shift-invariant point-spread function (PSF). By selecting proper physical constraints for optimization (e.g. choosing incident angles to be close together), the design of structures that transcend shift-invariance follows naturally. This is key for allowing our system to differentiate between angles in the paraxial regime. We emphasize that such differentiation is not possible under the assumption of shift-invariance, which is commonly used as a computational simplification [1], as is shown in Fig. 1(a,b). We numerically compute the measurement matrix G from $\varepsilon(\mathbf{r})$ using the finite-difference time-domain (FDTD) method [31].

2.2. Parameter estimation

Our pipeline is made up of a nanophotonic structure and a reconstruction algorithm. First, the electric fields generated on the detector by the ground truth object $\mathbf{u}$ are calculated with the image formation process, as described in the previous section. The result of that first step is the raw, noisy vector of intensities $\mathbf{v}$. We feed this vector into the computational reconstruction algorithm. The computational reconstruction algorithm uses elastic-net regression to reconstruct an object $\mathbf{u}_{\text{est}}$. Elastic-net regression is a form of linear regression with both a Lasso (l_1) and a Tikhonov (l_2) regularization term. Mathematically, the reconstruction problem amounts to solving the following optimization problem:

$$\mathbf{u}_{\text{est}} = \arg\min_{\mathbf{u}} \left(\|\mathbf{v} - G(\varepsilon(\mathbf{r}))\mathbf{u}\|_2^2 + \alpha_2 \|\mathbf{u}\|_2^2 + \alpha_1 \|\mathbf{u}\|_1 \right), \tag{2}$$

where α_1 and α_2 are the reconstruction hyperparameters. α_1 controls the magnitude of the Lasso (l_1) regularization term, which selects for sparsity. α_1 controls the magnitude of the Tikhonov (l_2) regularization term, which selects for a noise-robust solution. Our end-to-end optimization task is to find values for $\varepsilon(\mathbf{r})$, α_1, and α_2 that minimize the normalized reconstruction error averaged over the training set and corresponding image noise

$$L(\mathbf{u}_{\text{est}}, \mathbf{u}) = \left\langle \frac{\|\mathbf{u} - \mathbf{u}_{\text{est}}\|_2^2}{\|\mathbf{u}\|_2^2} \right\rangle_{\mathbf{u}, \eta} .$$

Our end-to-end optimization task to computationally design an imaging system can therefore be written mathematically as:

$$\varepsilon(\mathbf{r})^{(\text{opt})}, \alpha_1^{(\text{opt})}, \alpha_2^{(\text{opt})} = \underset{\varepsilon(\mathbf{r}), \alpha_1, \alpha_2}{\arg \min} \; L(\mathbf{u}, \mathbf{u}_{\text{est}}). \tag{3}$$

To perform optimization in the end-to-end framework, we need to compute gradients for the loss L with respect to the parameters $\varepsilon(\mathbf{r})$, α_1, and α_2. Gradients are back-propagated through the elastic-net reconstruction algorithm by finding the derivatives of the Karush-Kuhn-Tucker conditions [11], whereas back-propagation through the FDTD simulation is performed via an adjoint simulation. The adjoint simulation is run as a second FDTD [31] simulation following the initial "forward-pass" simulation. The adjoint sources are computed based on the back-propagated gradient vector from the reconstructed solution [32]. We then run FDTD using these adjoint sources and integrate the resulting fields with our forward-pass fields to obtain the overall gradient. We train our end-to-end system in a two-step process: first, we use the method of moving asymptotes (MMA) [33, 34] on a fixed training set for topology optimization of the nanophotonic structure $\varepsilon(\mathbf{r})$; then, we use Adam optimization [35] on data randomly generated each iteration to tune the reconstruction hyperparameters α_1 and α_2. For both MMA and Adam, we run the optimization until convergence; in practice, this occurred with a few hundred iterations of MMA and about 100 iterations of Adam. We benchmark our optimized designs against randomly initialized designs and thin lens designs by optimizing α_1 and α_2 for each design and comparing the optimized loss L. Because we randomly generate new data for each iteration of Adam, the training loss serves also as validation loss. To implement material binarization, we gradually turn on a binary filter for the structure (in practice, a tanh function) over the course of optimization. We begin optimization without a filter. After approximately fifty iterations of MMA, we start to filter the values of ε by the tanh function. Every approximately fifty iterations following, we turn up the sharpness of the filter, until the structure is binary.

Because the optimization is over such a large design space and the problem is generally non-convex, we expect the optimization to converge to local optima. Sometimes, edge cases of converging to local optima may be experienced where, for instance, the optimization converges to a structure that blocks all the light from the detector. To avoid these, we run our optimization several times from different initial starting conditions (uniform, random, random with Gaussian filter). For non-edge cases of convergence, however, the size of the design space means that the local optima we find generally perform very well for the task.

3. Results

We showcase our end-to-end pipeline in three types of reconstruction problems where, in order to accurately reconstruct the ground truth, we need finer angular resolution than what can be offered by paraxial optics. In all three reconstruction problems, we show angular resolutions that, in a paraxial system, would conventionally require over 20 times the separation we use between the lens and sensor (d), observing the measurement matrices breaking shift-invariance to resolve angles in the paraxial regime. In the overdetermined case, we also refer to this effect as "space

squeezing." We define the "compression ratio" as the factor by which we reduce the separation between the lens and sensor from the minimum required in paraxial imaging. In the first reconstruction problem (Fig. 2), we solve the 2D "sparse sensing" problem: an underdetermined inverse problem in two dimensions with a sparse prior, which means we assume there are far fewer nonzero elements in the ground truth than the total possible size of the ground truth; for instance, we may be detecting the angle and frequency of incoming laser beams, with the assumption that there are a relatively few number of incoming beams at any given time. The flexibility of our elastic-net reconstruction along with the requirement of sparsity allows us to accurately recover the ground truth signal. Over the course of optimization, our reconstruction algorithm settles into the compressed sensing limit of elastic-net, increasing the l_1 (Lasso) normalization coefficient and shrinking the l_2 (Tikhonov) normalization coefficient. In the second reconstruction problem (Fig. 3), we solve the 2D space squeezing problem: an overdetermined inverse problem in two dimensions with no sparsity prior. Here, the image reconstruction algorithm emphasizes the l_2 coefficient instead. In the third reconstruction problem (Fig. 4), we generalize the space squeezing problem to three dimensions.

Throughout the results, we report condition numbers of the measurement matrix $G(\varepsilon(\mathbf{r}))$, calculated as $\|G\|_2 \cdot \|G^+\|_2$, where G^+ is the pseudoinverse of G. Intuitively, a lower condition number means that the matrix is more robust to noise for reconstruction. Qualitatively, we find that the most reconstruction improvement for condition numbers that start in the range $10^2 - 10^3$ and decrease by a factor of 1.5 or more. For our thin lens benchmark, we can likewise model the optical system linearly as a measurement matrix and compute the condition number of that matrix. We evaluate our end-to-end imaging by root mean square error (RMSE) and structural similarity index measure (SSIM) [36] between ground truth and reconstructed signal. We calculate RMSE by the equation

$$\left\langle \frac{\|\mathbf{u} - \mathbf{u}_{\text{est}}\|_2^2}{\|\mathbf{u}\|_2^2} \right\rangle_{\mathbf{u}, \eta},$$

where $\mathbf{u}$ is the ground truth and $\mathbf{u}_{\text{est}}$ is the reconstructed signal.

We perform the 2D optimizations with 240 CPUs over the span of 2 to 3 days, and we perform the 3D optimizations with 480 CPUs over the span of 3 to 4 days. During topology optimization, we optimize the value of $\epsilon(\mathbf{r})$ continuously at each pixel, gradually turning on a binary filter over the course of optimization. For the 2D problems, we also gradually turn on a Gaussian filter to increase the feature size of the nanophotonic structure.

3.1. Two-dimensional sparse spectral-angular sensing

In this example, we show how end-to-end optimization can be used to reconstruct the spectrum and angle of incidence of an object beyond the paraxial limit for sparse ground-truth objects. Sparse sensing has application to laser awareness, enabling the simultaneous sensing of a small number of distinct signals from different directions and frequencies. The ground-truth object has dimensions 10 angles × 3 frequencies, making the vector representation of the ground-truth $\mathbf{u}$ a 30-component vector. The angles are uniformly spaced between −0.04 radians and 0.04 radians from normal incidence, such that the nonzero angles for each ground-truth $\mathbf{u}$ are drawn from this particular set of 10 angles. The frequencies correspond to red (672 nm), green (560 nm), and blue light (448 nm). Each detector pixel has length $x_D = 3.36\,\mu m$. We set the sensor $d = 11.2\,\mu m$ from the structure. In the sparse sensing problem, we optimize the nanophotonic structure in a $78\,\mu m \times 1\,\mu m$ design region. We first demonstrate our method in a 2D setting. Here we use a sensor with 20 detector pixels, which is 2/3 times the total number of parameters to be reconstructed. From the specifications above, this makes the entire sensor $67.2\,\mu m$ long. We give our nanophotonic structure a design region of size $78\,\mu m \times 1\,\mu m$ (Fig. 2(a)), with 1296
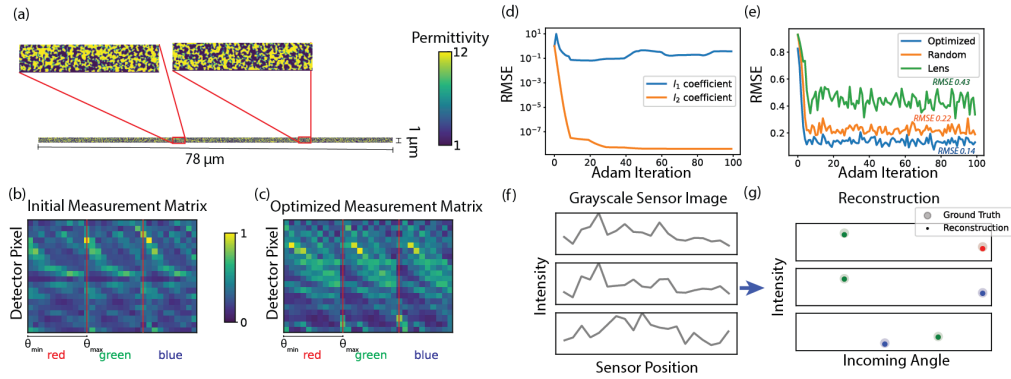
Fig. 2. **Sparse, underdetermined angle-resolved spectrometry.** (a) The optimized $78\,\mu\mathrm{m} \times 1\,\mu\mathrm{m}$ volumetric nanophotonic structure, represented as a binary heatmap of $\varepsilon(\mathbf{r})$, with zoom-ins on two regions of the structure. There are 1296 design degrees of freedom per $1\,\mu\mathrm{m}^2$. In the final optimized structure, $\varepsilon(\mathbf{r})$ takes on permittivities of only 1 and 12. (b, c) Measurement matrix G (b) before and (c) after optimization, with detector pixel position on the y-axis and incoming angle and color on the x-axis. (d) Evolution of the l_1 (Lasso) and l_2 (Tikhonov) regularization coefficients over optimization. (e) Convergence of the reconstruction error during end-to-end optimization (blue) compared to convergence of the reconstruction error during reconstruction-only optimization with a random structure (orange) and convergence of the reconstruction error during reconstruction-only optimization with a thin lens (green). (f) Grayscale images formed from the ground truths in Fig. 2(g) on the detector with 1% noise. (g) Sparse ground truth signal of 2 activated elements out of a total of 10 angle × 3 frequency possibilities, overlaid with reconstructed sparse signal from the image in Fig. 2(f).

degrees of freedom per $1\,\mu\mathrm{m}^2$.

We initialize the nanophotonic structure as a random binary structure. This results in the measurement matrix shown in Fig. 2(b). The measurement matrix is constructed by propagating monochromatic plane waves that span the length of the nanophotonic structure and measuring the result on the sensor for each plane wave. The measurement matrix is then indexed by sensor pixel along the rows and by incoming angle and frequencies along the columns. The vertical red lines separate the matrix into sections by incoming plane wave frequency; within each section, the incoming angles range over all 10 angles. Because the sensor is smaller than the structure and therefore smaller than the full span of the incoming plane waves, the entire sensor initially detects near-constant low intensity over all incoming angles and frequencies.

After optimization, using the procedure described in the previous section, the measurement matrix is shown in Fig. 2(c) and the optimized structure in Fig. 2(a). Qualitatively, the optimized matrix has less correlation between adjacent pixels than the initial measurement matrix, which matches findings that matrices with independent random pixels are well-suited for compressed sensing problems [37]. The sensitivity (or robustness) of the reconstruction to environmental noise can be characterized using the condition number of the measurement matrix, defined as the ratio of maximal to minimal singular value. Quantitatively, over the course of the optimization, the condition number of the measurement matrix decreases from 2431 to 1514, and the optical transmission increases from 0.21 to 0.27 (Table 1). We note that the columns of the optimized measurement matrix (Fig. 2(c)) differ substantially from each other, meaning the imaging system is not shift-invariant in the angles. This property is key for accurately resolving adjacent angles in the paraxial regime. By comparison, the columns for the measurement matrix of a thin

lens are the same across different incoming angles. Meanwhile, on the reconstruction side, the optimization emphasizes the l_1 regularization coefficient while making the l_2 coefficient shrink to a negligible value, nearly five orders of magnitude under the l_1 coefficient (Fig. 2(d)). Here, the presence of the l_2 regularization term improves convergence in the early iterations of reconstruction hyperparameter optimization even though it is eventually set to nearly zero, as observed in [11]. Altogether, by giving the end-to-end optimization an underdetermined sparse problem, the optimization computationally settles on both Lasso regression and a nanophotonic structure that leads to a randomized measurement matrix, conditions consistent with existing compressed sensing literature. We emphasize that we specified neither of those conditions as explicit goals of our optimization.

Overall, the optimized system takes in the input signal from a sparse, multichromatic ground truth, forming a noisy, randomized, grayscale image (Fig. 2f), and accurately recovering the ground truth by solving the compressed sensing problem with Lasso regression (Fig. 2g). The optimized end-to-end system accurately recovers incoming sparse signals under 1% sensor noise with RMSE 0.14, a significant improvement over a random structure (RMSE 0.22) and a thin lens focusing to the detector (RMSE 0.43) (Fig. 2e).

From the parameters described previously, the interval between adjacent angles is $\Delta\theta = \frac{\pi}{360}$. This puts us well beyond the paraxial limit, as $\tan(\Delta\theta) \approx 0.0087 < 0.30 = \frac{x_D}{d}$. In terms of space squeezing, this gives us an effective compression ratio $\frac{x_D/d}{\tan(\Delta\theta)}$ of 34.5 over an angular bandwidth of 0.08 radians and a spectral bandwidth of $2.2 \cdot 10^{14}$ Hz. For a comparison of angular and spectral ranges, see Table 2. Because we are beyond the paraxial limit, this is a situation where a traditional lens would do poorly—in particular, multiple adjacent angles often give the same reading on the sensor with a traditional lens, illustrated in Fig. 1.

3.2. Two-dimensional polychromatic space squeezing

This second reconstruction problem shares many of the same design parameters as the other two-dimensional problem described in section 3.1 (same ground-truth vector dimensions, angle range, frequencies, and detector parameters). However, in the case of general-purpose space squeezing, one cannot *a priori* assume sparsity of the ground truth image.

We therefore demonstrate our method in an overdetermined system without the sparse prior. Here we activate all 30 components of the 10 angles $\times$ 3 frequencies ground-truth object. We use a sensor with 50 pixels, so our reconstruction algorithm is solving an overdetermined regression problem. From the specifications, the overall sensor is then $168\,\mu$m long. Our nanophotonic design region is extended to $134\,\mu$m $\times 1\,\mu$m to better match the sensor size (Fig. 3(a)), again with 1296 degrees of freedom per $1\,\mu$m^2. . We again initialize the nanophotonic structure as a random structure, leading to the measurement matrix Fig. 3(b) and the structure depicted in Fig. 3(a). After optimization, the measurement matrix increases significantly in maximum intensity and is no longer shift-invariant (Fig. 3(b,c)). Quantitatively, over the optimization, the condition number decreases from 499 to 166, and the transmission increases from 0.20 to 0.53 (Table 1, Fig. 3(h)). Here sparsity is not enforced in the ground-truth, so the regularization coefficients no longer emphasize the l_1 term in the hyperparameter tuning (Fig. 3(d)). An example ground truth, noisy grayscale image, and reconstructed signal are shown in Fig. 3(f,g). The optimized system has structural similarity index measure, or SSIM, 0.86 and RMSE 0.10 at 1% Gaussian image noise, a significant improvement over a random structure (SSIM 0.64, RMSE 0.18) and a lens (SSIM 0.42, RMSE 0.25).

To go from the sparse underdetermined problem to the general overdetermined problem, we only had to change the physical conditions of the pipeline. In particular, we made no changes to the initial reconstruction hyperparameters, with the optimization automatically choosing to emphasize the l_1 regularization term and reduce the l_2 term to nearly zero, the opposite of what we had previously observed in Fig. 2(d). The flexibility of elastic-net and end-to-end
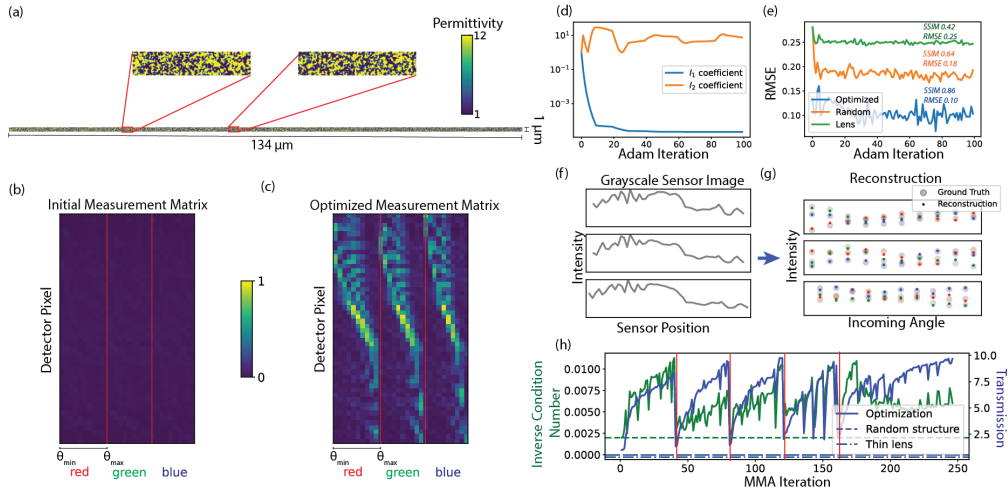
Fig. 3. **General angle-resolved spectrometry on a large sensor.** (a) The optimized $134\,\mu\text{m} \times 1\,\mu\text{m}$ volumetric nanophotonic structure, represented as a binary heatmap of $\varepsilon(\mathbf{r})$, with zoom-ins on two regions of the structure. There are 1296 design degrees of freedom per $1\,\mu\text{m}^2$. In the final optimized structure, $\varepsilon(\mathbf{r})$ takes on binary permittivities of only 1 and 12. (b) The measurement matrices G (b) before and (c) after optimization. (d) Evolution of the l_1 (Lasso) and l_2 (Tikhonov) regularization coefficients over optimization. (e) Convergence of the reconstruction error during end-to-end optimization. (f) Grayscale images formed from the ground truths in Fig. 3(g) on the detector with 1% noise. (g) Ground truth signal of 10 angles × 3 frequencies overlaid with reconstructed sparse signal from the image in Fig. 3(f). (h) Comparison of inverse condition numbers and transmissions between optimized structure, random structure, and thin lens. Red vertical lines indicate new training phases, increasing binary filter strength on the structure with each new phase.

optimization allows us to solve these different classes of problems without any manual tuning of the reconstruction algorithm. We also emphasize the application of the general overdetermined problem for space squeezing, or imaging in systems with compact free space.

Here, $\frac{x_D}{d} = 0.30$ and the compression ratio is 34.5, same as in the previous example.

3.3. Three-dimensional monochromatic space squeezing

Our last example reconstruction problem is a 3D extension of the result in Section 3.2. Here, like in the 2D space squeezing problem, the reconstruction problem is overdetermined. In this scenario, we set our ground-truth object to be of only one frequency and 5 x-angles × 5 y-angles, so the ground-truth vector u is a length-25 vector. The angles here are spaced between -0.02 radians and 0.02 radians from normal incidence in both the $x-$ and $y-$ dimensions. Each detector pixel is a square with side length $x_D = 2.24\,\mu m$. We set the sensor $d = 11.2\,\mu\text{m}$ from the structure. We allow all 25 components of the ground-truth to be activated in the optimization. We use an 11×11 sensor, which makes the overall sensor have size $24.6\,\mu\text{m} \times 24.6\,\mu\text{m}$, The nanophotonic design region is of size $13.4\,\mu\text{m} \times 13.4\,\mu\text{m} \times 0.56\,\mu\text{m}$ (Fig. 4(a)), with 10648 design degrees of freedom per $1\,\mu\text{m}^3$. Here, we perform freeform optimization over voxels, as opposed to alternative of optimizing 2D patterns.

Before optimization, we initialize the nanophotonic as a random structure, which forms the measurement matrix shown in Fig. 4(b). After optimization, the measurement matrix becomes the one shown in Fig. 4(c) and the nanophotonic design becomes the structure shown in
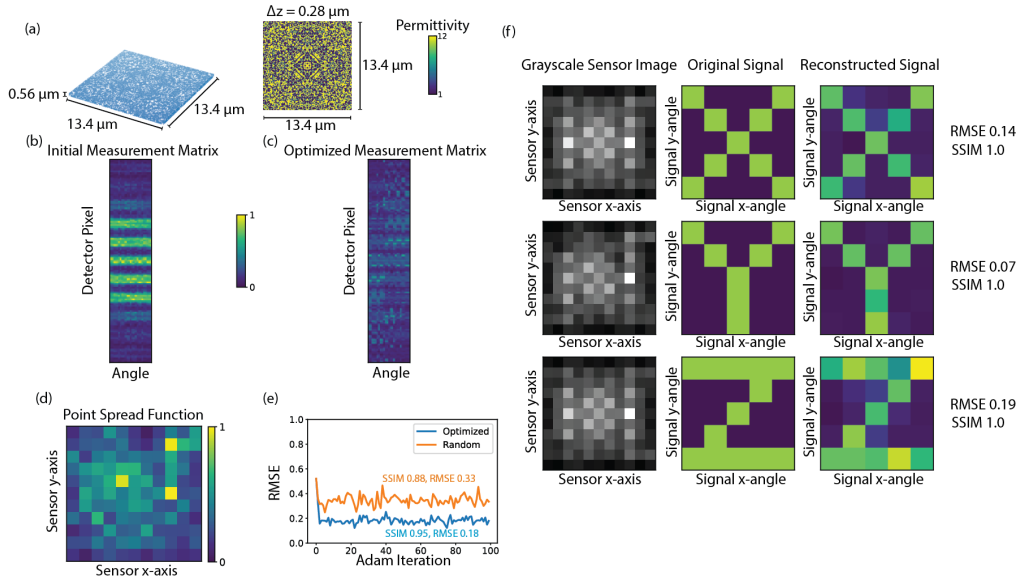
Fig. 4. **3D Imager.** (a) Cross-sections of the optimized $13.4\,\mu\text{m} \times 13.4\,\mu\text{m} \times 0.56\,\mu\text{m}$ volumetric nanophotonic structure at $\Delta z = 0.28\,\mu\text{m}$ from the surface of the nanophotonic structure, with 10648 design degrees of freedom per $1\,\mu\text{m}^3$. In the final optimized structure, $\varepsilon(\mathbf{r})$ takes on binary permittivities of only 1 and 12. (b, c) The measurement matrices G (b) before and (c) after optimization. (d) Point spread function measured for the normal-incidence original signal. (e) Convergence of the reconstruction error during end-to-end optimization. (f) Side-by-side examples of grayscale sensor image, original signal, and reconstructed signal.

Fig. 4(a). Qualitatively, the optimized measurement matrix is more sparse, and different angles are more focused on the sensor than in the initial measurement matrix, which serves to make the optimized matrix better-conditioned. Quantitatively, the condition number of the measurement matrix decreases from 902 to 221 over the optimization (Table 1). However, unlike in the 2D reconstruction problems, the transmission slightly decreases here, likely a result of the limited design space. We also show the point spread function for the optimized structure (Fig. 4(e)). The intensity in the point spread function is localized in the top left corner. In the optimized measurement matrix (Fig. 4(c)), the point spread function corresponds to the middle column, and the high intensity can be seen at the top of the row. Three example reconstructions are shown in Fig. 4(f); qualitatively the reconstructed signals generally faithfully capture the main features of the original signals. Overall, the optimized system is benchmarked to have SSIM 0.95 and RMSE 0.18 at 1% noise, a significant improvement over the system with a random structure (SSIM 0.88, RMSE 0.33).

Lastly, this 3D structure achieves $\frac{x_D}{d} = 0.20$ and its compression ratio is 22.9, beyond the realm of paraxial optics.

4. Discussion and outlook

Our central contribution is a flexible, noise-robust framework for transcending shift-invariance while imaging in the paraxial regime. By designing a volumetric nanophotonic structure with topology optimization for our optical element, we are no longer beholden to the shift-invariant paraxial approximation. Compared to traditional lenses or conventional metasurfaces (relying on the locally periodic approximation), we can keep the detector closer and the sensor resolution

	Condition Number	Transmission
2D thin lens, sparse	2.36×10^7	0.90
2D random structure, sparse	2431	0.21
2D optimized structure, sparse	1514	0.27
2D thin lens, space squeezing	1.65×10^{10}	0.90
2D random structure, space squeezing	499	0.20
2D optimized structure, space squeezing	166	0.53
3D random structure, space squeezing	902	0.39
3D optimized structure, space squeezing	221	0.36

Table 1. Condition numbers and transmissions over various designs.

	Compression Ratio	Maximum Angle (radians)	Wavelength Range (nm)
Reshef et al., metamaterial spaceplate	4.9	0.26	1550
Reshef et al., uniaxial spaceplate	1.12	0.61	visible light (400–700)
Guo et al., three-layer	144	0.01	dependent on design constant (single wavelength)
Guo et al., single-layer hexagonal	11.2	0.11	dependent on design constant (single wavelength)
Our work, 2D (sparse and space squeezing)	34.5	0.04	460–690
Our work, 3D (space squeezing)	22.9	0.02	550

Table 2. Comparison of our work with other space squeezing designs.

lower, which lets us keep the entire imaging system compact. We demonstrate that our method significantly outperforms both a thin lens and a random scattering structure in paraxial imaging with compression ratios of up to 34.5. In comparison, previous works have demonstrated compression factors of up to 4.9 [18] and 144 [2]. The combination of computation and structure design is the strength of our work, showing enhanced performance compared to lensless imaging (computation only) or a photonic inverse-designed structure (photonics only). The addition of a post processing step does not present a significant hurdle to real-life applications, given the relative simplicity of our reconstruction algorithm (compared to state-of-the-art artificial intelligence tasks implemented on centralized or edge computing platforms [38, 39]). We note that previous work in building spaceplates [2, 18] results in a general optical element that directly implements the transfer function of free space. On the other hand, our work builds an optical element that is combined with a reconstruction algorithm to result in a compact imaging system. Indeed, these two approaches are not in conflict with each other; future work could involve using a spaceplate in conjunction with a freeform nanophotonic structure to further reduce the free space required in the design. Other approaches in improving imaging resolution have been proposed, including Fourier ptychography [40] and multi-aperture, folded-optics imaging [41]. Fourier ptyochgraphy focuses on the trade-off between resolution and field-of-view, whereas we focus on the trade-off between resolution and free-space. Multi-aperture imaging addresses the resolution to depth tradeoff as we do, but it focuses on building new designs for the camera aperture and folding the imaging system to achieve a longer equivalent free space within the same space envelope. On the other hand, we design a new optical element to replace the lens component of a camera, not changing the aperture or folding the entire system.

We noted above in our examples (summarized in Table 1) that our optimization either preserves

or improves transmission by the nanophotonic structure. For instance, the two-dimensional space squeezing optimization shows a 2.5x increase in transmission. This leads to a twofold benefit—the improved reconstruction accuracy shows that the system becomes more noise-robust, and the increase in transmission shows the system increases the signal-to-noise ratio.

Future work may build on our framework by innovating on either the nanophotonic design or reconstruction algorithm. The choice of freeform nanophotonics opens our design to many more possibilities than prior works with locally-periodic metasurfaces, but there is a tradeoff in computational cost. For instance, to optimize a freeform nanophotonic structure in three dimensions (Section 3.3), we had to significantly reduce our problem size. With more computational power or more efficient design choices, the nanophotonic design could be scaled up to explore richer physics and higher-resolution imaging. For instance, this could be done with Flexcompute or by imposing an axisymmetric structure to reduce computational costs [42]. Innovations on the reconstruction algorithm may include replacing our elastic-net reconstruction with more general algorithms, such as neural networks. To push the transmission of our inverse-designed structures closer to that of a real lens, future work may focus on performing end-to-end optimization with transmission encoded as part of the loss function, for instance adding in a term to the loss function to penalize low transmission. This work would be important for improving the transmission of optimized structures to allow them to be considered for real applications.

To experimentally realize our designs, there are additional fabrication constraints to account for during topology optimization, such as minimum length scales and connectivity [43, 44]. We do not presently account for these constraints in this proof-of-concept work. Our work is theoretical in nature, but we note that existing volumetric fabrication technologies (such as implosion fabrication) would make an experimental demonstration of our method possible in the near future [45, 46]. Additionally, we could modify our optimization method to account for already-demonstated nanofabrication techniques, such as multilayer metasurfaces [47–49].

Looking forward, we anticipate a growing demand for compact imaging. Our end-to-end framework coupled with freeform nanophotonics paves the way for the design of optical elements that can perform high-resolution imaging with limited volume. In particular, we present our method as a more general alternative to optics-only space squeezers. It is our hope that the application of end-to-end design to compact imaging will allow for smaller, higher-resolution cameras.

Disclosures. The authors declare no conflicts of interest.

Data availability. The data underlying the results presented in this paper are not publicly available at this time but may be obtained from the authors upon reasonable request.

References

1. J. W. Goodman, "Introduction to fourier optics, 3rd ed." CO: Roberts & Co. Publ. **1** (2005).
2. C. Guo, H. Wang, and S. Fan, "Squeeze free space with nonlocal flat optics," Optica **7**, 1133–1138 (2020).
3. N. Yu and F. Capasso, "Flat optics with designer metasurfaces," Nat. materials **13**, 139–150 (2014).
4. A. Arbabi, Y. Horie, M. Bagheri, and A. Faraon, "Dielectric metasurfaces for complete control of phase and polarization with subwavelength spatial resolution and high transmission," Nat. nanotechnology **10**, 937–943 (2015).
5. M. Khorasaninejad, W. T. Chen, R. C. Devlin, J. Oh, A. Y. Zhu, and F. Capasso, "Metalenses at visible wavelengths: Diffraction-limited focusing and subwavelength resolution imaging," Science **352**, 1190–1194 (2016).
6. H.-T. Chen, A. J. Taylor, and N. Yu, "A review of metasurfaces: physics and applications," Reports on Prog. Phys. **79**, 076401 (2016).
7. J. Engelberg and U. Levy, "The advantages of metalenses over diffractive lenses," Nat. communications **11**, 1–4 (2020).
8. J. S. Jensen and O. Sigmund, "Topology optimization for nano-photonics," Laser & Photonics Rev. **5**, 308–321 (2011).
9. R. E. Christiansen and O. Sigmund, "Inverse design in photonics by topology optimization: tutorial," JOSA B **38**, 496–509 (2021).

10. S. Molesky, Z. Lin, A. Y. Piggott, W. Jin, J. Vucković, and A. W. Rodriguez, "Inverse design in nanophotonics," Nat. Photonics **12**, 659–670 (2018).
11. G. Arya, W. F. Li, C. Roques-Carmes, M. Soljačić, S. G. Johnson, and Z. Lin, "End-to-end optimization of metasurfaces for imaging with compressed sensing," arXiv preprint arXiv:2201.12348 (2022).
12. V. Sitzmann, S. Diamond, Y. Peng, X. Dun, S. Boyd, W. Heidrich, F. Heide, and G. Wetzstein, "End-to-end optimization of optics and image processing for achromatic extended depth of field and super-resolution imaging," ACM Trans. on Graph. (TOG) **37**, 1–13 (2018).
13. Z. Lin, C. Roques-Carmes, R. Pestourie, M. Soljačić, A. Majumdar, and S. G. Johnson, "End-to-end nanophotonic inverse design for imaging and polarimetry," Nanophotonics **10**, 1177–1187 (2021).
14. Z. Lin, R. Pestourie, C. Roques-Carmes, Z. Li, F. Capasso, M. Soljačić, and S. G. Johnson, "End-to-end metasurface inverse design for single-shot multi-channel imaging," Opt. Express **30**, 28358–28370 (2022).
15. Q. Sun, J. Zhang, X. Dun, B. Ghanem, Y. Peng, and W. Heidrich, "End-to-end learned, optically coded super-resolution spad camera," ACM Trans. on Graph. (TOG) **39**, 1–14 (2020).
16. E. Tseng, A. Mosleh, F. Mannan, K. St-Arnaud, A. Sharma, Y. Peng, A. Braun, D. Nowrouzezahrai, J.-F. Lalonde, and F. Heide, "Differentiable compound optics and processing pipeline optimization for end-to-end camera design," ACM Trans. on Graph. (TOG) **40**, 1–19 (2021).
17. C. Spägele, M. Tamagnone, D. Kazakov, M. Ossiander, M. Piccardo, and F. Capasso, "Multifunctional wide-angle optics and lasing based on supercell metasurfaces," Nat. Commun. **12**, 1–10 (2021).
18. O. Reshef, M. P. DelMastro, K. K. Bearne, A. H. Alhulaymi, L. Giner, R. W. Boyd, and J. S. Lundeen, "An optic to replace space and its application towards ultra-thin imaging systems," Nat. communications **12**, 1–8 (2021).
19. A. Overvig and A. Alù, "Diffractive nonlocal metasurfaces," Laser & Photonics Rev. **16**, 2100633 (2022).
20. K. Shastri and F. Monticone, "Nonlocal flat optics," Nat. Photonics pp. 1–12 (2022).
21. D. A. Miller, "Why optics needs thickness," Science **379**, 41–45 (2023).
22. F. Monticone, "Toward ultrathin optics," Science **379**, 30–31 (2023).
23. K. Shastri, O. Reshef, R. W. Boyd, J. S. Lundeen, and F. Monticone, "To what extent can space be compressed? bandwidth limits of spaceplates," Optica **9**, 738–745 (2022).
24. Z. Lin, C. Roques-Carmes, R. E. Christiansen, M. Soljačić, and S. G. Johnson, "Computational inverse design for ultra-compact single-piece metalenses free of chromatic and angular aberration," Appl. Phys. Lett. **118**, 041104 (2021).
25. K. Yanny, N. Antipa, W. Liberti, S. Dehaeck, K. Monakhova, F. L. Liu, K. Shen, R. Ng, and L. Waller, "Miniscope3d: optimized single-shot miniature 3d fluorescence microscopy," Light. Sci. & Appl. **9**, 171 (2020).
26. G. Satat, M. Tancik, and R. Raskar, "Lensless imaging with compressive ultrafast sensing," IEEE Trans. on Comput. Imaging **3**, 398–407 (2017).
27. K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th on International Conference on Machine Learning,* (Omnipress, Madison, WI, USA, 2010), ICML'10, p. 399–406.
28. J. Zhang and B. Ghanem, "Ista-net: Interpretable optimization-inspired deep network for image compressive sensing," in *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition,* (2018), pp. 1828–1837.
29. E. Markley, F. L. Liu, M. Kellman, N. Antipa, and L. Waller, "Physics-based learned diffuser for single-shot 3d imaging," in *NeurIPS 2021 Workshop on Deep Learning and Inverse Problems,* (2021).
30. Q. Bertrand, Q. Klopfenstein, M. Blondel, S. Vaiter, A. Gramfort, and J. Salmon, "Implicit differentiation of lasso-type models for hyperparameter optimization," in *Proceedings of the 37th International Conference on Machine Learning,* (2020).
31. A. F. Oskooi, D. Roundy, M. Ibanescu, P. Bermel, J. D. Joannopoulos, and S. G. Johnson, "Meep: A flexible free-software package for electromagnetic simulations by the fdtd method," Comput. Phys. Commun. **181**, 687–702 (2010).
32. S. G. Johnson, "Notes on adjoint methods for 18.335," Introd. to Numer. Methods (2012).
33. S. G. Johnson, "The NLopt nonlinear-optimization package," (2017).
34. K. Svanberg, "A class of globally convergent optimization methods based on conservative convex separable approximations," SIAM journal on optimization **12**, 555–573 (2002).
35. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv preprint arXiv:1412.6980 (2014).
36. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," IEEE transactions on image processing **13**, 600–612 (2004).
37. E. J. Candès and M. B. Wakin, "An introduction to compressive sampling," IEEE signal processing magazine **25**, 21–30 (2008).
38. P. Raina, "Energy-efficient circuits and systems for computational imaging and vision on mobile devices," Ph.D. thesis, Massachusetts Institute of Technology (2018).
39. J. Chen and X. Ran, "Deep learning with edge computing: A review," Proc. IEEE **107**, 1655–1674 (2019).
40. G. Zheng, C. Shen, S. Jiang, P. Song, and C. Yang, "Concept, implementations and applications of fourier ptychography," Nat. Rev. Phys. **3**, 207–223 (2021).
41. G. Carles and A. R. Harvey, "Multi-aperture imaging for flat cameras," Opt. Lett. **45**, 6182–6185 (2020).
42. R. E. Christiansen, Z. Lin, C. Roques-Carmes, Y. Salamin, S. E. Kooi, J. D. Joannopoulos, M. Soljačić, and S. G. Johnson, "Fullwave maxwell inverse design of axisymmetric, tunable, and multi-scale multi-wavelength metalenses," Opt. Express **28**, 33854–33868 (2020).

43. A. M. Hammond, A. Oskooi, M. Chen, Z. Lin, S. G. Johnson, and S. E. Ralph, "High-performance hybrid time/frequency-domain topology optimization for large-scale photonics inverse design," Opt. Express **30**, 4467–4491 (2022).
44. A. M. Hammond, A. Oskooi, S. G. Johnson, and S. E. Ralph, "Photonic topology optimization with semiconductor-foundry design-rule constraints," Opt. Express **29**, 23916–23938 (2021).
45. D. Oran, S. G. Rodriques, R. Gao, S. Asano, M. A. Skylar-Scott, F. Chen, P. W. Tillberg, A. H. Marblestone, and E. S. Boyden, "3d nanofabrication by volumetric deposition and controlled shrinkage of patterned scaffolds," Science **362**, 1281–1285 (2018).
46. F. Han, S. Gu, A. Klimas, N. Zhao, Y. Zhao, and S.-C. Chen, "Three-dimensional nanofabrication via ultrafast laser patterning and kinetically regulated material assembly," Science **378**, 1325–1331 (2022).
47. M. Mansouree, H. Kwon, E. Arbabi, A. McClung, A. Faraon, and A. Arbabi, "Multifunctional 2.5 d metastructures enabled by adjoint optimization," Optica **7**, 77–84 (2020).
48. P. Camayd-Muñoz, C. Ballew, G. Roberts, and A. Faraon, "Multifunctional volumetric meta-optics for color and polarization image sensors," Optica **7**, 280–283 (2020).
49. C. Roques-Carmes, Z. Lin, R. E. Christiansen, Y. Salamin, S. E. Kooi, J. D. Joannopoulos, S. G. Johnson, and M. Soljačić, "Toward 3D-Printed Inverse-Designed Metaoptics," ACS Photonics **9**, 43–51 (2022).