Supporting Information for:

Array-Level Inverse Design of Beam Steering Active Metasurfaces

Prachi Thureja[1],[†], Ghazaleh Kafaie Shirmanesh[1],[†], Katherine T. Fountaine[2], Ruzan Sokhoyan[1], Meir Grajower[1], and Harry A. Atwater[1],*

[1]Thomas J. Watson Laboratory of Applied Physics, California Institute of Technology, Pasadena, California 91125, USA

[2]NG Next, Northrop Grumman Corporation, One Space Park, Redondo Beach, CA 90278, USA

[†]P. Thureja and G. Kafaie Shirmanesh contributed equally to this work.

[*]E-mail: haa@caltech.edu

KEYWORDS: inverse design, active metasurface, phased array, beam steering, array configuration, genetic algorithm, wavefront engineering

1. Independent scatterer model for subwavelength antenna arrays

The electric far-field calculation, as introduced in equation (1) of the main manuscript, is based on antenna array theory that was initially developed for phased arrays operating at radio frequencies (RF).[1] As RF phased arrays can maintain large interantenna spacings without the introduction of additional diffraction orders, their far-field response can be modelled as that of a structured collection of independent scattering elements. However, this assumption becomes non-trivial in the case of metasurfaces with subwavelength antenna spacings operating at visible or near-infrared wavelengths, as the characteristic period d_x between neighboring antennas approaches the electromagnetic near-field regime given by $2d_x^2/\lambda$.[1] Thus, coupling between neighboring elements can only be neglected for antennas that are spaced at large enough distances and/or possess strongly confined modes. While the electromagnetic near-field of the indium tin oxide (ITO) based metasurface studied in this work was extensively examined by Kafaie Shirmanesh *et al.*,[2] we demonstrate here an alternative approach to verify the validity of the independent scatterer model.

Interantenna coupling causes the actual array phase and amplitude profile to deviate from the target design. Consequently, it manifests itself as an increase in the relative magnitude of the sidelobes due to undesired interference effects. While near-field coupling is not accounted for in the analytical model, it can be observed in full-wave simulations of the entire metasurface configured with the target array profile. Thus, we perform a comparison of the analytically calculated radiation patterns to the full-wave simulations[2] for forward-designed array profiles (Fig. S1). We would like to bring the reader's attention to two apparent effects that are observed in the cases shown in Fig. S1. Firstly, rather than an increase in the relative magnitude of the sidelobes (that would indicate interantenna coupling), Fig. S1a-c depict an attenuation of the sidelobes at broadside angles in the simulated radiation patterns. This effect increases the simulated directivity D_{sim} in comparison to its analytically computed counterpart D_{AF} that is based on array factor calculations. Secondly, a strong increase in the relative magnitude of the zero-order sidelobe and thus decrease in D_{sim} is reported for $\theta_r = 70.7°$ (Fig. S1d). To understand the observed deviations, we remind ourselves that the analytical calculations were performed for omnidirectional scatterers with $E_{antenna} = 1$. While this assumption is valid for a broad range of steering angles with the current nanoantenna design it breaks at larger angles. Thus, the reported deviations are attributed to a simplified antenna model rather than a manifestation of interantenna coupling.
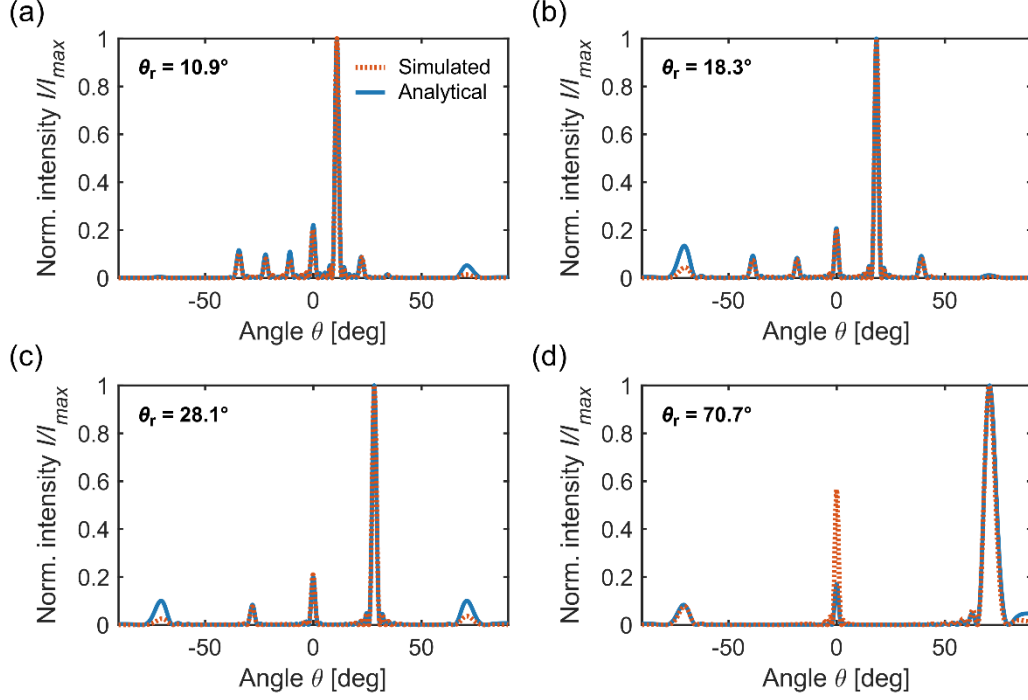
Figure S1. Normalized far-field radiation patterns I/I_{max} as a function of the polar angle obtained through full-wave simulations[2] (orange dashed) and analytical array factor calculations (blue). The results are obtained using forward-designed stairstep array profiles for an array of 96 antennas arranged at a period of d_x = 400 nm. The operating wavelength is λ = 1510 nm. The simulated (D_{sim}) and analytically calculated (D_{AF}) directivities at the respective steering angles θ_r are: (a) D_{sim} = 51.0 and D_{AF} = 43.5, (b) D_{sim} = 48.3 and D_{AF} = 39.5, (c) D_{sim} = 50.1 and D_{AF} = 39.1, (d) D_{sim} = 22.2, D_{AF} = 23.0.

2. Phase gradient profiles

Forward-designed phase gradient profiles rely on a constant phase shift φ_s between adjacent antennas. For an incident beam normal to the array, φ_s is computed as[3,4]

$$\varphi_s = 360° \cdot \frac{d_x \cdot \sin(\theta_r)}{\lambda}. \tag{S1}$$

Here, d_x is the characteristic period between neighboring phase antennas, θ_r is the steering angle, and λ is the operating wavelength. Wrapping of the phase profiles around 360° allows the design of blazed grating-like structures that steer the reflected beam in the desired direction. However, due to a discrete sampling of the phases at fixed spatial increments d_x, the blazed grating of an antenna array comprises of discontinuous steps, as shown in Fig. S2a. The discrete sampling further results in an aperiodicity of the phase profiles (Fig. S2b), as a complete phase shift of 360° is not necessarily an integer multiple of φ_s. The phase profiles approach periodic blazed gratings for all steering angles as d_x goes to smaller values.
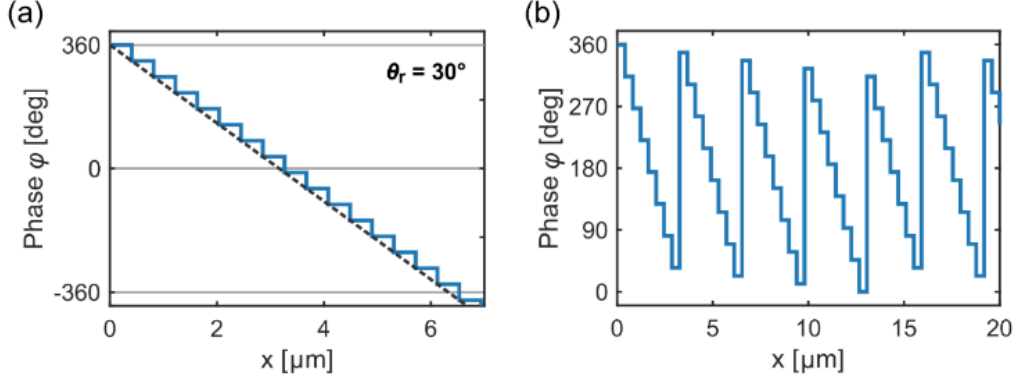
Figure S2. (a) Phase φ as a function of the spatial x-coordinate in an antenna array. The black dashed line represents an unwrapped, continuous phase profile for steering at $\theta_r = 30°$. The blue curve illustrates the corresponding discretization steps due to sampling at constant spatial increments $d_x = 400$ nm. The horizontal grey lines mark the edge values for wrapping around $360°$. The operating wavelength is $\lambda = 1550$ nm. (b) Phase φ vs. spatial x-coordinate for discretized phase profile wrapped around $360°$. The aperture is extended to 20 µm to display the aperiodicity of the wrapped phase profile.

3. Beam steering performance metrics: directivity *vs.* power efficiency

The figure of merit (FOM) quantifying the beam steering performance in this work was chosen to be the beam directivity D. It is a unitless quantity that depends on the ratio of the intensity at the desired steering angle θ_r to the amount of power scattered into all directions normalized by the solid angle, as discussed in equation (2) of the manuscript. Thus, it remains unaffected by scaling of the far-field radiation patterns by a constant factor. Directivity is a common metric used to analyze the performance of RF phased arrays. An ideal metasurface array with $d_x = 400$ nm operating at $\lambda = 1550$ nm ($d_x/\lambda \sim 0.25$) approaches performances that are reported with an array of parallel short dipoles.[5] In addition, the optimized sidelobe level reported in this work corresponds to values that are generally obtained for phased arrays with a complete phase modulation over $360°$.[6]

The power efficiency η is determined by the absolute amount of power that is steered into the main lobe compared to the total input power. For an array profile with varying amplitudes, η is calculated as

$$\eta(\theta_r) = \frac{P_m(\theta_r)}{P_{scat}} \cdot A_{eq} \tag{S2}$$

where P_m is the power scattered into the main lobe steering at θ_r and P_{scat} is the total scattered power. The ratio of P_m and P_{scat} is multiplied by the equivalent amplitude A_{eq} that would be required in an array of antennas with constant amplitude to generate an equivalent amount of total scattered power. Thus, $A_{eq} = P_{scat}/P_{input}$ with P_{input} being the input power. Note that P_{input} can be determined by assuming an ideal reflectarray with constant, unit amplitude and a complete phase modulation over $360°$.

Due to the strong absorption in the active antenna element, the power efficiency of the beam steering arrays studied in this work[2] is strongly limited. Consequently, the optimized directivity case discussed in the manuscript (Fig. S3a-b) results a power efficiency of 0.9%, even though 86% of the total scattered power is directed into the main lobe. Here, we demonstrate as a proof-of-concept that the same inverse design algorithm can also be applied to a power efficiency optimization. For this purpose, the figure of merit is adapted to FOM = $\eta(\theta_r)$. Figure S3c shows the optimized array profile as well as the corresponding radiation pattern (Figure S3d) for optimal power efficiency at $\theta_r = 18.3°$. It is to be noted that the increase in power efficiency comes at the cost of beam directivity, as the algorithm aims to increase the occurrence of large amplitudes in the antenna array to enhance efficiency. Therefore, the amplitude modulation increases, leading to a reduction in beam directivity. Meanwhile, the opposite trend holds true for a directivity optimization: The inverse design aims to minimize amplitude modulation to reduce sidelobes. As the main phase shift occurs in a low amplitude regime, the minimization of amplitude modulation results in reduced power efficiencies. For reference, the corresponding directivity and efficiency values are tabulated in Table S1.
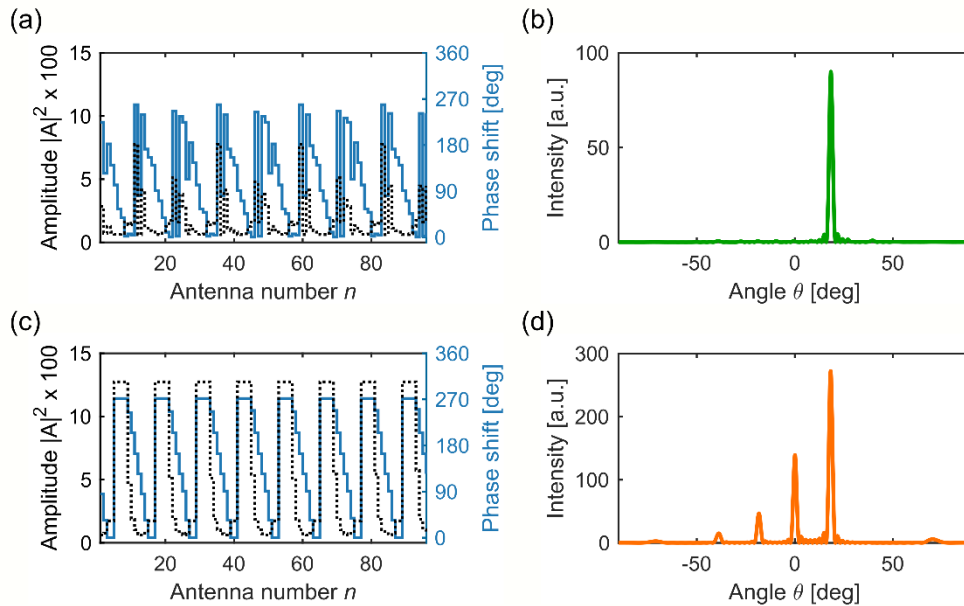

Figure S3. Inverse-designed array amplitude (black dotted) and phase (blue) profiles and far-field radiation patterns for an optimization of directivity (a)-(b) and an optimization of power efficiency (c)-(d), respectively. Optimizations are performed for the ITO-based metasurface studied in this work.[2] The corresponding directivity and power efficiency values are listed in Table S1.

	Directivity D	Efficiency η
Forward design, stairstep	39.5	2.1%
Inverse design, directivity opt.	72.7	0.9%
Inverse design, efficiency opt.	41.9	2.7%

Table S1. Directivity D and efficiency η, respectively, for the electro-optically tunable, ITO-based metasurface introduced by Kafaie Shirmanesh et al.[2] The corresponding values are compared for three different cases: forward-designed stairstep array profiles, and inverse-designed profiles optimized for directivity and efficiency, respectively.

As the scattered light amplitudes are the limiting factor for power efficiencies in beam steering metasurfaces, we would like to remark that they can be strongly enhanced with the use of active metasurfaces exhibiting higher reflectance / transmittance values, such as all-dielectric metasurfaces.[7-9]

4. Complex dielectric permittivity of indium tin oxide (ITO)

The continuity of the normal electric displacement component at the interface between two media requires

$$\varepsilon_1 \cdot E_{1\perp} = \varepsilon_2 \cdot E_{2\perp} \tag{S3}$$

where ε_i is the complex dielectric permittivity and $E_{i\perp}$ $E_{i\perp}$ is the normal electric field component in medium i. Hence, operation at an epsilon-near-zero (ENZ) condition results in a strong field enhancement in the active ITO layer. The spectral overlap of this ENZ transition with the magnetic dipole resonance of the antenna thus ensures a strong modulation of the scattered light response.

The spatial variation of the real and imaginary parts of the ITO permittivity (ε_{ITO}) under applied bias are presented in Figure S4. The ITO properties are chosen as described in Ref. 10. As can be seen in Fig. S4a, for a sufficiently large applied bias, an ENZ condition holds in the ITO layer where the real part of the ITO permittivity can take values between -1 and +1. Figure S4b shows that the imaginary part of the ITO permittivity in the mentioned regions takes nonzero values. The nonzero complex permittivity of the ITO layer in the ENZ region leads to a finite electric field confinement in the accumulation region of the ITO. Nevertheless, owing to the near-zero real part of the permittivity, the ITO layer experiences a large field enhancement, as shown in the electromagnetic near-field distributions provided by Kafaie Shirmanesh et al.[2] in Parts 2 and 3 of the Supporting Information of their manuscript.
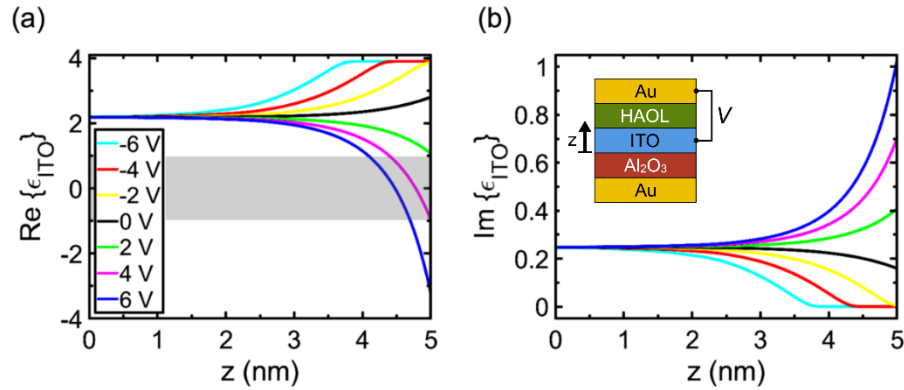

Figure S4. Spatial variation of complex dielectric permittivity of ITO under applied bias. (a) Real and (b) imaginary part of the permittivity of ITO as a function of position for different applied bias voltages at the operating wavelength of $\lambda = 1510$ nm. The gray-shaded region in (a) shows the ENZ regime where $-1 < \text{Re}\{\varepsilon_{ITO}\} < +1$. $z = 5$ nm denotes the interface of the ITO and the HAOL gate dielectric, and $z = 0$ nm represents the interface of the ITO and the Al$_2$O$_3$ dielectric layer, as indicated in the inset of (b). Voltage is applied between the ITO and the top Au layer.
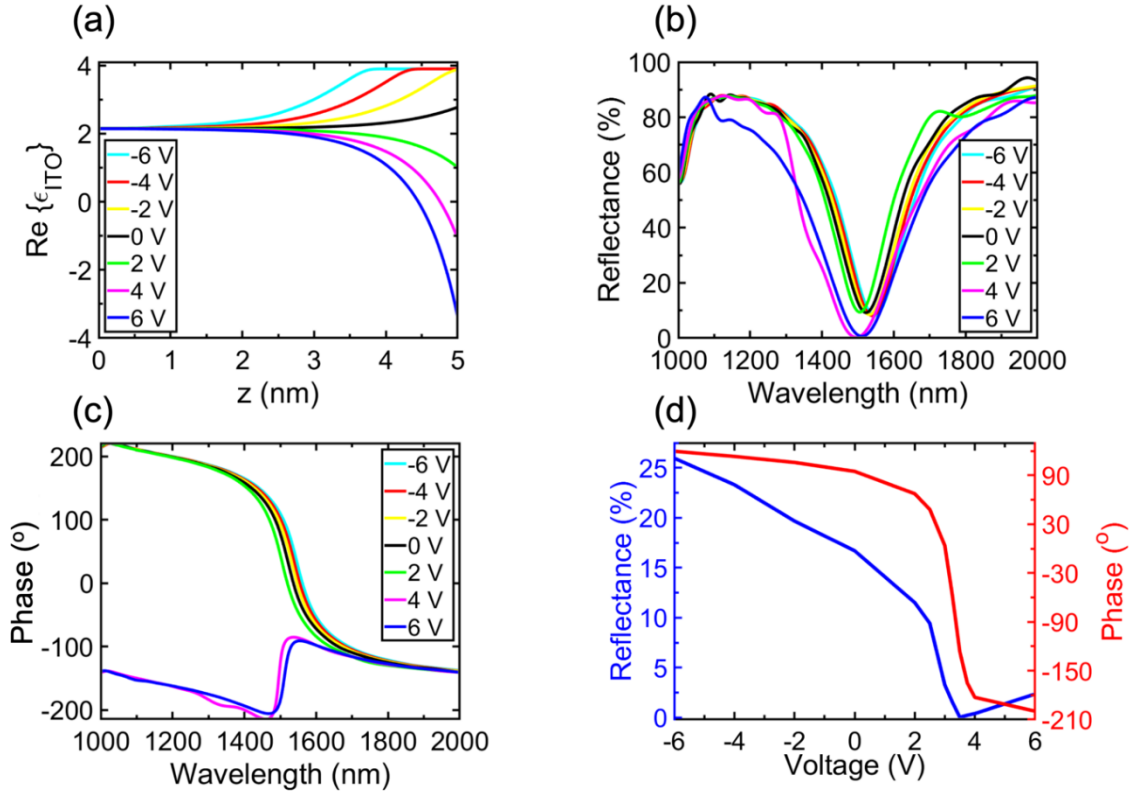
Figure S5. Optical response of the metasurface incorporating an artificial ITO layer with zero collision frequency, and hence zero imaginary part of permittivity. (a) Spatial variation of the real part of the permittivity of the artificial ITO film as a function of position for different applied bias voltages at the wavelength of $\lambda = 1510$ nm. Here, $z = 5$ nm denotes the interface of the ITO and the HAOL gate dielectric, and $z = 0$ nm represents the interface of the ITO and the Al_2O_3 dielectric layer. (b) Reflectance and (c) phase of reflection spectra for different applied biases. (d) Reflectance and phase of the reflection as a function of applied bias at the operating wavelength of $\lambda = 1490$ nm.

According to the Drude model, the complex permittivity of the ITO can be formulated as

$$\varepsilon_{ITO}(\omega) = \varepsilon_\infty - \frac{\omega_p^2}{\omega^2 + i\omega\Gamma} \tag{S4}$$

where ε_∞ is the infinite frequency permittivity, ω is the angular frequency, ω_p is the plasma frequency and Γ is the collision frequency. The latter contributes to Im(ε_{ITO}) and quantifies the losses in form of absorption through electron-electron collisions. Thus, for increased values of Im(ε_{ITO}), lower scattered light amplitudes are expected as a result of enhanced absorption.[11] Here, we simulate the same metasurface structure with an artificial ITO layer with zero imaginary part. To this end, we set the collision frequency of the ITO layer to be zero ($\Gamma = 0$ in eq. (S4)). Figure S5a presents the real part of the permittivity of the artificial ITO layer as a function of position (with $z = 5$ nm being the interface of the ITO and the HAOL gate dielectric) for different applied bias voltages. For sufficiently large applied bias voltages, an ENZ condition holds in the artificial

ITO layer. The spectra of the reflection amplitude (reflectance) and phase are depicted in Fig. 5b and c, respectively. Upon changing the applied bias, a notable reflectance and phase modulation could be obtained. Figure S5d illustrates the reflectance and phase values as a function of applied bias at the operating wavelength of λ = 1490 nm. As can be seen, a remarkable amplitude modulation with increased amplitude values at negative voltages, accompanied by a phase shift of 320° can be achieved *via* incorporating the mentioned artificial ITO layer within our metasurface. Notably, reflectance is reduced at ENZ condition even though the active layer is lossless. This implies that Au absorbs more light due to a stronger field confinement in the gap. Furthermore, the resonant position has been shifted due to the zero imaginary part and a slight change in the real part of the permittivity of the artificial ITO layer compared to the original ITO film. Finally, we would like to note that while there have been theoretical studies on active metasurfaces that are based on low-loss materials with near-zero complex dielectric permittivity, such as cadmium oxide (CdO),[12] they have yet to be demonstrated experimentally.

5.   Forward designs in non-ideal antenna arrays

Device non-idealities for active metasurfaces include tunable antennas with (i) non-unity amplitudes, (ii) reduced phase modulation range, and (iii) covarying phase and amplitude values. While non-unity amplitudes impact the power efficiencies of steered beams, the latter two device characteristics are directly translated into limitations of forward designs, as discussed below. The phase and amplitude profiles shown in Fig. S6 are obtained with the scattered light properties of the electro-optically tunable metasurface in Ref. 2.

Reduced phase modulation range

A reduced phase modulation range in an active metasurface requires modification to the phase gradient profile to ensure that it remains within the maximal acquired phase shift. Here, we discuss two possible adjustments to ideal forward designs. Linear truncated phase profiles consist of blazed gratings that are truncated symmetrically around 180°. Phase profiles are then shifted such that the minimal acquired phase value is 0° (Fig. S6a, blue). Step profiles, on the other hand, simplify the design of phase gradient profiles by repeating a discrete number of phase values over several antenna (Fig. S6c, blue). As the truncated linear phase profiles have a closer resemblance to ideal blazed gratings, higher directivities are reported in that case.

Covariation of amplitude and phase

Covarying amplitude and phase values inhibit the design of pure phase gratings, as indicated by the black dotted lines in Fig. S6a, c. As a consequence, increased destructive interference results in additional scattering that appears in form of undesired sidelobes. Figure S6b, d illustrate how the far-field radiation patterns change after consideration of the phase-amplitude interdependencies for steering at θ_r = 18.3°. For the two cases considered here, directivities drop from $D_{lin,const}$ = 72.1 to $D_{lin,covary}$ = 50.7 for linear phase profiles, and from $D_{step,const}$ = 54.8 to $D_{step,covary}$ = 39.5 for stairstep phase profiles.
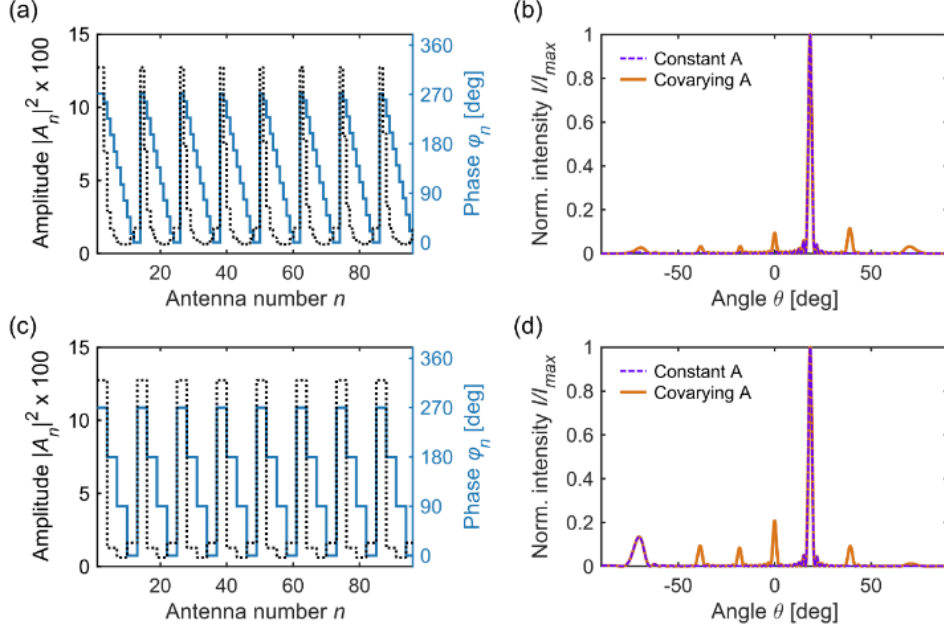
Figure S6. (a) Linear truncated phase (blue) and corresponding amplitude profile (black dotted) over 96 antennas for steering at $\theta_r = 18.3°$. (b) Normalized far-field intensity I/I_{max} *vs*. polar angle θ for linear truncated phase profiles with constant amplitudes (violet dashed) and covarying amplitudes (orange). (c) Stairstep phase (blue) and corresponding amplitude profile (black dotted) over 96 antennas for steering at $\theta_r = 18.3°$. The stairstep phase profile is obtained by periodically repeating each phase shift of 270°, 180°, 90°, 0° over 3 consecutive antennas. (d) Normalized far-field intensity I/I_{max} *vs*. polar angle θ for stairstep phase profiles with constant amplitudes (violet dashed) and covarying amplitudes (orange). The phase-amplitude relation is obtained from the optical response of the electro-optically tunable metasurface introduced in Ref. 2. Antennas are arranged at a period of $d_x = 400$ nm. The operation wavelength is $\lambda = 1510$ nm. Radiation patterns are computed with the assumption of omnidirectional antennas.

6. Iterative genetic optimization: numerical framework

Figure S7 outlines the implementation of iterative genetic algorithms using the global optimization toolbox on MATLAB. The input of the algorithm comprises of the steering angle θ_r, as well as the objective function $FOM(x, \varphi(V), A(V))$ that takes into account the tunable scattered light properties of the metasurface. x is the 1D vector representing the array configuration that needs to be optimized. In addition, we define the following global variables: the total number of antennas N_{tot}, the number of optimization rounds r_{tot}, as well as an array containing the number of possible variables *nvars* which are to be optimized in each iteration. For the active metasurface with 96 tunable antennas, *nvars* is defined as [4, 8, 24, 48, 96] such that the optimal solution is found within a maximal number of five iterations. The concept of iterative genetic optimization relies on an initially reduced search space.
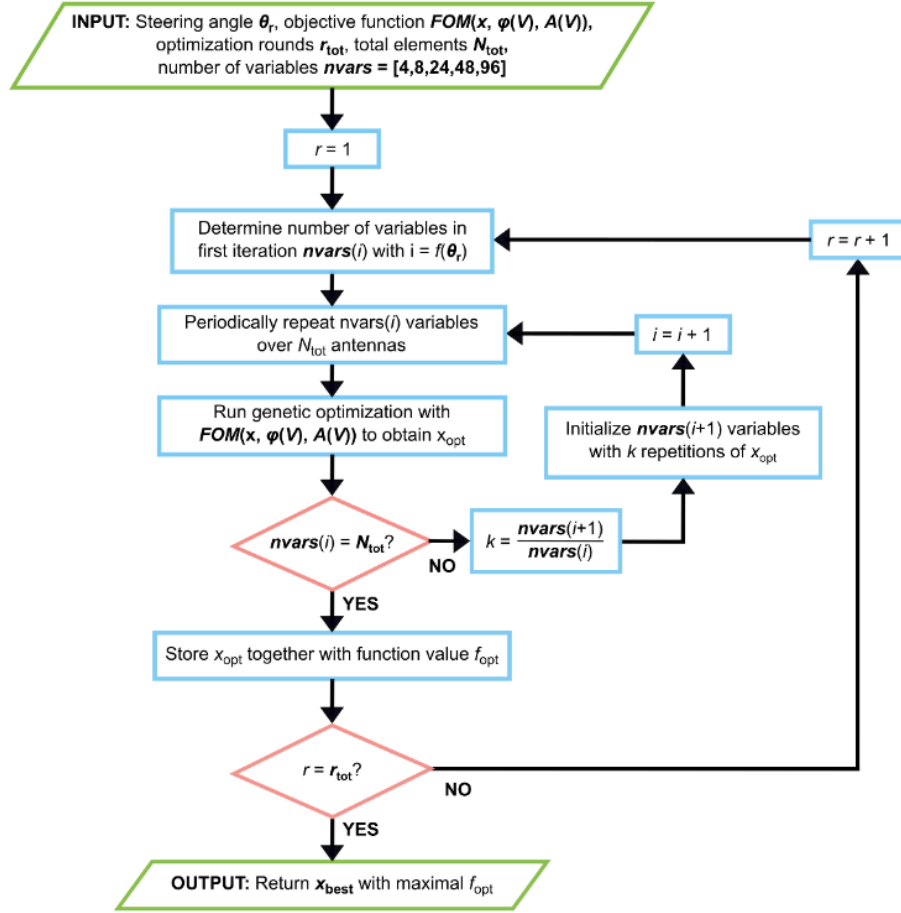
Figure S7. Flowchart of iterative genetic optimization for an array of $N_{tot} = 96$ antennas. The inner loop represents the iterative genetic optimization with increasing variable size to approach the high dimensionality of the underlying problem. The outer loop describes a series of optimization rounds that allow to take the optimal solution over multiple repetitions. The latter is required due to the stochastic nature of genetic optimization.

The algorithm aims to optimize for a sequence of small number of variables that are periodically repeated over the entire array. Once an optimal solution x_{opt} is found, *nvars* is incrementally increased to the next value. An initialization with $k = nvars(i+1)/nvars(i)$ repetitions of the current optimized solution x_{opt} guides the algorithm in larger solution domains. This procedure is repeated until all N_{tot} antennas are considered as free variables in the final iteration. Once $nvars(i) = N_{tot}$, the current optimization round is terminated and x_{opt} is stored along with its corresponding function value in an array. This iterative optimization process is repeated for r_{tot} rounds, after which the solution with the maximal f_{opt} is given as output. This step is necessary due to the stochastic nature of genetic optimization. Note that prior knowledge from blazed grating design allows us to make the algorithm more efficient. The number of variables that are to be optimized in the first iteration can be determined as a function of the steering angle θ_r, using the grating equation defined in (1).

Convergence properties

The iterative optimization process for the active metasurface analyzed in our study is illustrated in the supporting movie 'MovieS1.avi'. The target steering angle is $\theta_r = 18.3°$. As can be seen in the convergence plot, the main contribution to the increase in directivity occurs in the initial optimization in the reduced solution space. Once an optimal solution is found, the algorithm moves on to the next bigger solution domain where minor changes in phase and amplitude lead to a reduction of the sidelobes which are displayed in the log-scaled inset of the far-field radiation pattern as well as in the best directivity value. Since a separate optimization was performed in order to collect the required data at each generation, the optimized directivity differs from the value reported in the manuscript.

Computational cost

In contrast to forward-designed array profiles that rely on an analytical equation, as discussed in (S1), the inverse design approach comes with enhanced computational cost due to a consideration of the antenna-specific functional response. For the problems analyzed in our work, the optimal solution in each iteration is generally obtained within 200-600 generations. A single computation ($r_{tot} = 1$) for an iterative optimization of 96 variables performed on our workgroup computer (Intel Xeon E5-2687W processor, 20 cores) takes approximately 12 min.

The required computation time highly depends on the total number of variables that are to be optimized. Figure S8 shows the average computation time TOC_{avg} for a single iterative optimization round as a function of N_{tot}. The average time was evaluated over $r_{tot} = 10$ optimization rounds for six different steering angles ($\theta_r = 9.0°$, 10.9°, 13.6°, 18.3°, 28.1°, 70.7°). Notably, the computation time scales linearly as $O(N_{tot})$ in the investigated regimes while the solution space scales exponentially as $O(s^{N_{tot}})$ where s is the number of sampling points. For our study, the antenna-specific scattered light response was sampled at $s = 65$ discrete voltage points. The difference in scaling is attributed to the stopping criteria: In the current implementation, the algorithm stops once the average change in best function value over 250 generations is less than 10^{-6}. As the most significant contribution of the directivity enhancement occurs for the initial optimization in a reduced solution domain, each subsequent iteration adds approximately 250 generations to the optimization process that result in minor performance enhancements. Therefore, a linear increase in computation time is observed. This phenomenon can also be seen in the supporting movie 'MovieS1.avi', illustrating the convergence of the iterative optimization. In future work, the stopping criteria can be optimized such that the computational cost is reduced without a significant loss in best performance.
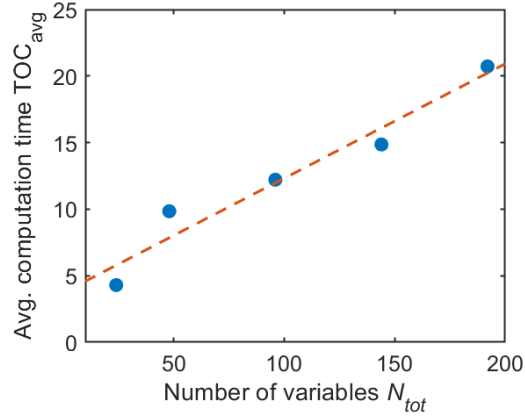
Figure S8. Average computation time TOC_{avg} as a function of N_{tot} variables that are to be optimized. Data points TOC_{avg} (blue) are obtained as a mean over $r_{tot} = 10$ optimization rounds and six different steering angles ($\theta_r = 9.0°$, $10.9°$, $13.6°$, $18.3°$, $28.1°$, $70.7°$). The linear relation between TOC_{avg} and N_{tot} is illustrated with an orange dashed line.

7. Comparison to alternative optimization methods in terms of robustness

Genetic algorithms rely on a random initial population that contains possible solutions to a given problem. Selection procedures only permit survival of the best solutions to the next generation. Operators inspired by natural genetic variation (crossover and mutation) further introduce variability into population members. Due to the stochastic operations in genetic optimization, convergence characteristics differ between individual optimization rounds.[13] Thus, it is common practice to report the optimal solution as the one with maximal *FOM* over an extended dataset obtained over r_{tot} optimization rounds. To verify the robustness of the best result, we analyzed the distribution of the optimized directivity over $r_{tot} = 20$ optimization rounds. In addition, we perform a comparison of the distribution to two alternative optimization methods: a direct, non-iterative optimization of the entire antenna array with an initial guess based on linear phase profiles and a direct optimization without a user-defined initial guess. In the latter case, the algorithm generates a random initial solution to seed the algorithm. Figure S9a shows the range of optimized directivities for the three analyzed methods for a steering angle of $\theta_r = 18.3°$. In all three cases the beam directivity is strongly enhanced in comparison to forward designs. Direct optimization of 96 variables with an initial guess based on forward-designed linear phase profiles results in a maximal increase in directivity of 77% compared to the previously demonstrated stairstep forward design with $D_{forward} = 39.5$. Meanwhile, an increase of up to 80% is reported with a direct optimization using a randomly generated initial guess. In comparison, the iterative optimization approach which relies on an incremental increase of the solution space facilitates a maximal increase in directivity of up to 84%, as reported in the manuscript. While the optimized directivity approaches similar values in all three cases, there is a distinct difference in the robustness of the final result. The direct optimization with an initial guess based on forward design drives the algorithm to similar local minima, as the forward designs do not account for the antenna-specific amplitude-phase correlation (Fig. S9b). A subset of the solutions that can escape these local minima results in marginally higher directivities. By contrast, the direct optimization with a random initial guess (Fig. S9c) leads to stronger directivity enhancements due to an unbiased and

thus more extensive exploration space. Finally, the iterative approach proposed in this work relies on an optimization of the array profile in a reduced solution domain before passing the optimized result from the prior iteration as an initial guess to the next iteration. By doing so, this method ensures that the antenna-specific scattered light response is accounted for when supplying the algorithm with an initial solution in each iteration. As a result, higher directivities are obtained with increased robustness, as illustrated by the strongly increased median in the corresponding boxplot in Fig. S9a. For the case studied here, 75% of the optimized performances lie within 3% of the maximal directivity $D_{\mathrm{opt,iter}} = 72.7$ (Fig. S9d).
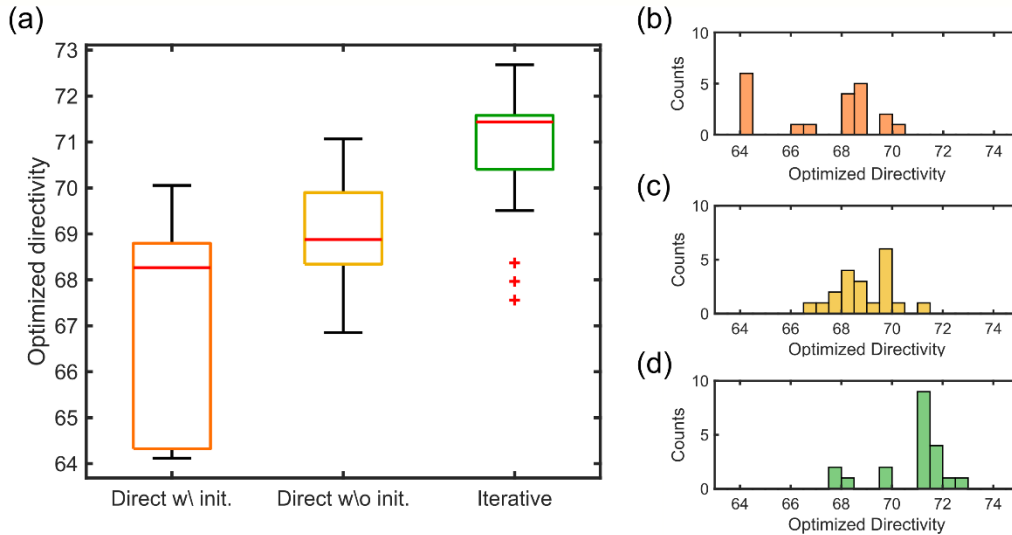

Figure S9. (a) Distribution of the optimized directivity over $r_{\mathrm{tot}} = 20$ optimization rounds for a steering angle of $\theta_{\mathrm{r}} = 18.3°$. The results are illustrated for three different optimization methods: a direct, non-iterative optimization approach of the entire array with an initial guess based on forward-designed linear phase profiles (orange), a direct optimization without a user-defined initial guess (yellow), and an iterative optimization approach with an incremental increase of the solution space (green). The red horizontal line in the boxplot marks the median of the distribution, while the upper and lower edges of the box indicate the 25th and 75th percentile. The whiskers extend to the most extreme optimized directivity points that are not considered outliers and are marked in red crosses. Histograms of the optimized directivity distributions with a bin width of 0.5 are illustrated in (b), (c), and (d), respectively.

8. Directivity and aperture size

The two main factors defining the beam directivity are the magnitude of the sidelobes relative to the peak intensity at the desired steering angle θ_{r}, as well as the beam divergence (*i.e.*, full width at half maximum FWHM) of the main lobe. While the prior is optimized in the inverse design process, the latter is determined based on the aperture size. Here, we illustrate that directivity can be enhanced by enlarging the aperture through an increase in the number of antennas while keeping the pitch fixed. In particular, we analyze the directivity of ideal antenna arrays at oblique angles where reduced values are generally reported due to the diminished effect of aperture size. As the

13

number of antennas is increased from 50 to 1000, a reduced FWHM (Fig. S10a) and thus increased directivity (Fig. S10b) is reported.
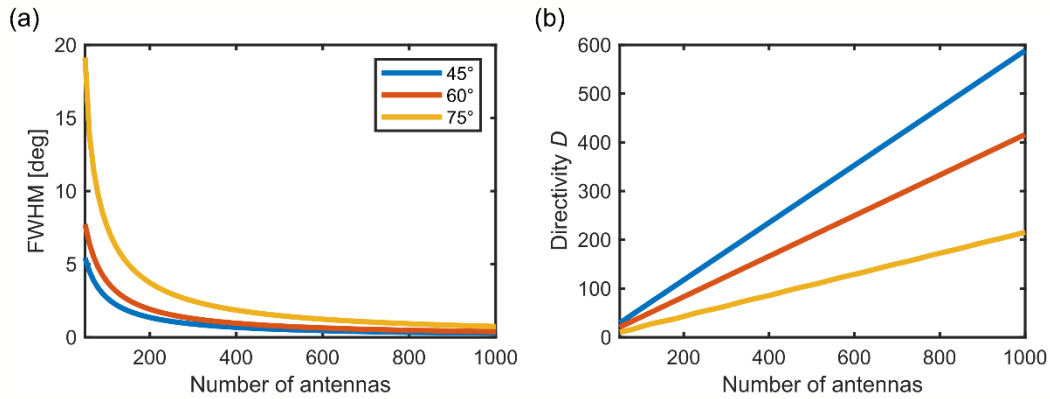

Figure S10. (a) FWHM and (b) directivity D for an ideal antenna array with a fixed pitch of $d_x = 400$ nm. The number of antennas is varied from 50 to 1000, corresponding to a change in aperture size from 20 μm to 400 μm. Results are illustrated for an operating wavelength of $\lambda = 1510$ nm.

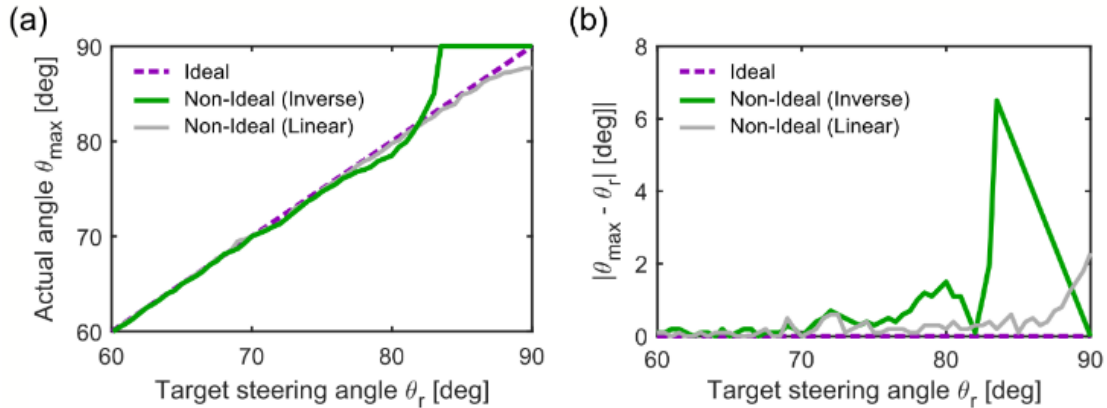## 9. Target *vs.* actual steering angle for continuous beam steering

Figure S11. (a) Actual angle θ_{max} *vs.* target steering angles θ_r for broadside angles. (b) Absolute deviation from target angle $|\theta_{max} - \theta_r|$ *vs.* target steering angle θ_r for broadside angles. Results are shown for forward-designed ideal phased array (violet dashed), as well as forward (grey) and inverse designs (green) for the non-ideal active metasurface.

Since steering at broadside angles requires large phase gradients, reduced phase modulation ranges limit the amount of information that can be carried by each slope. This limits the angular steering accuracy at broadside steering angles. Furthermore, optimization of the beam directivity aims to maximize the intensity at the desired steering angle θ_r, while minimizing the beam divergence as well as noise in form of sidelobes. Thus, by reducing the FWHM of the steered beam (Fig. 3b), inverse design can enhance beam directivity at broadside angles even when the actual steering angle θ_{max} does not correspond to θ_r. Figure S11a and b illustrate how θ_{max} and the absolute deviation from the desired steering angle $|\theta_{max} - \theta_r|$ evolve for forward- and inverse-

designed arrays. Depending on the desired application, steering accuracies can be improved by either defining an alternative objective function or by implementing additional constraints that set an upper barrier to the maximum permissible deviation from θ_r.

10. Phase modulations for analysis of hypothetical devices

To decouple the effects of phase and amplitude on the optimized performance, Fig. 5 outlines a study for a series of hypothetical devices. Here, the phase is modelled as a sigmoidal function of the applied bias. This behavior is typical for active metasurface that exhibit the largest phase shift at resonance conditions. To ensure consistent degrees of freedom in each case, acquired phase shifts are modelled over the same bias range. Figure S12 illustrates the corresponding phase characteristics for phase modulation ranges from 360° to 90°.
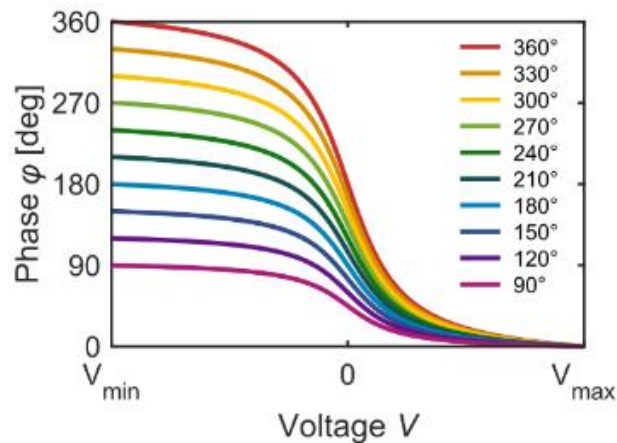

Figure S12. Acquired phase shift φ as a function of the applied bias voltage V for phase modulation ranges from 360° to 90°. Phase is assumed to be a sigmoidal function of the applied bias.

11. Optimized results for Lorentzian amplitude relation

Here, we discuss the inverse-designed structures obtained with a Lorentzian amplitude-voltage relation and a phase modulation of 360°, as outlined in Section 3 of the manuscript. Figure S13a illustrates the non-intuitive optimized amplitude and phase profiles for steering at $\theta_r = 15°$. Inverse design enables an increase in the directivity from $D_{\text{forward,lorentz}} = 66.5$ to $D_{\text{inverse,lorentz}} = 73.8$ by suppressing sidelobes, as shown in Fig. S13b.
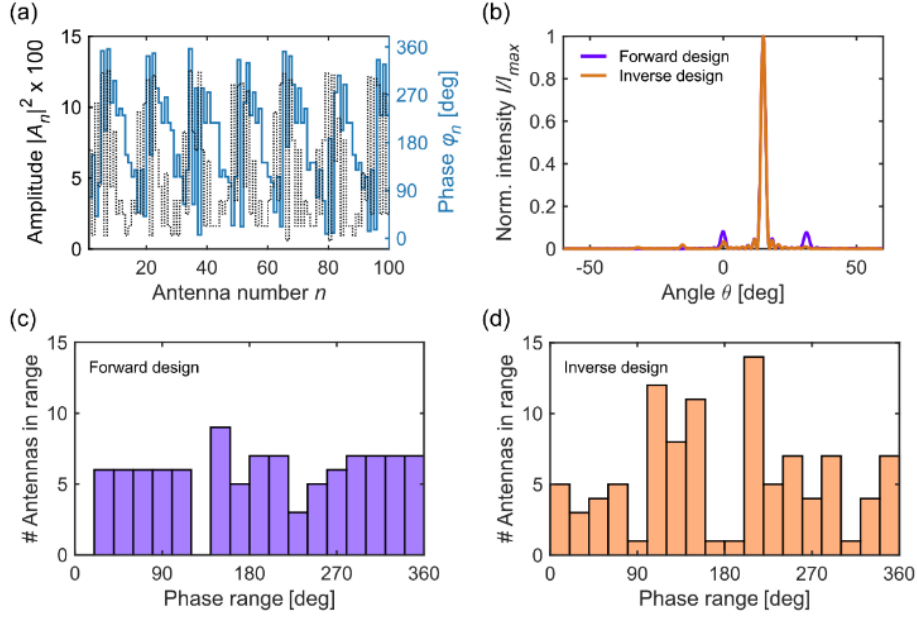
Figure S13. (a) Optimized phase and amplitude profile over 100 antennas for Lorentzian amplitude-voltage relation and a phase modulation of 360°. (b) Normalized far-field intensity I/I_{max} as a function of the polar angle θ for forward (linear truncated, violet) and inverse-designed (orange) arrays. The corresponding directivities are $D_{forward,lorentz} = 66.5$ and $D_{inverse,lorentz} = 73.8$. The difference in performance is supported by a change in phase distribution over 360°. The number of antennas in each phase range between 0° and 360° is displayed in (c) for forward design and in (d) for the inverse-designed phase profiles. The histogram bin width is 20°.

This difference is supported by altered distributions of the phase values over 360°. Since phase gradient profiles are based on constant phase shifts between adjacent antennas, the distribution of phases over the entire array is nearly uniform (Fig. S13c). Minor differences are attributed to a finite aperture size as well as the specific value of the phase shift that amounts to $\varphi_s = 22.3°$ in this case. Inverse design, on the other hand, aims to suppress sidelobes by minimizing the amplitude modulation. This can be achieved by avoiding low amplitude regimes. In the case of the Lorentzian amplitude-voltage relation, minimal amplitude is reported for an acquired phase of 180° (Fig. 5c). Therefore, inverse design results in a considerably smaller number of antennas with acquired phases in that regime (Fig. S13d).

12. Full-wave simulation of experimentally demonstrated metasurface

Figure S14 shows the full-wave simulation results of the optical response of the metasurface using finite difference time domain (FDTD) simulations. As can be seen in Fig. S14a, when changing the applied bias, reflectance modulation can be observed especially at bias voltages greater than 2 V. This reflectance modulation is accompanied by a significant phase shift, as shown in Fig. S14b. Figure S14c shows the achievable phase shift as a function of applied bias voltage at different wavelengths. A phase shift of 280° withing a bias voltage ranging from -6 V to +6 V is obtained at a wavelength of $\lambda = 1545$ nm which is selected as the operating wavelength of the beam steering metasurface.
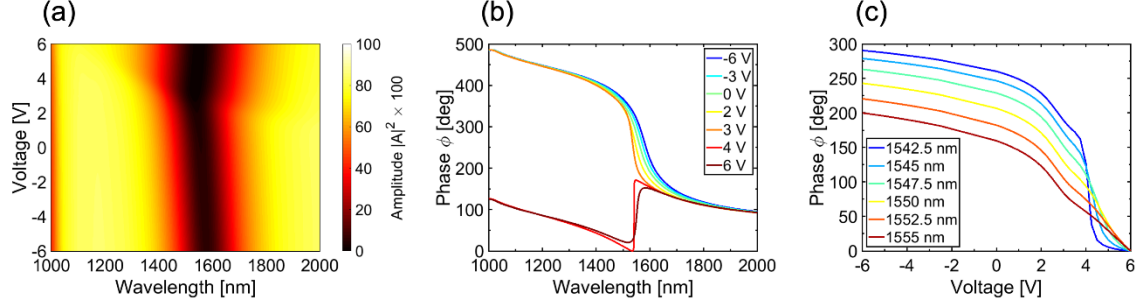
Figure S14. Full wave simulation results of (a) reflectance as a function of wavelength and applied bias voltage (b) acquired phase spectrum at different applied biases and (c) acquired phase as a function of applied bias at different wavelengths.

13. Experimental setup for phase and amplitude measurements

To characterize the tunable optical response of the fabricated beam steering metasurface, we measured the spectra of the reflected light amplitude (reflectance) and phase under different applied biases. Figure S15 shows the optical setup used to measure the phase shift as well as the reflectance modulation provided by the metasurface. In order to measure the phase of the light reflected from the metasurface, the metasurface sample is illuminated by a tunable near-infrared (NIR) laser which is focused on the sample by an objective with a long working distance (Mitutoyo M Plan Apo 20×, NA = 0.40, WD = 20 mm) after passing through a polarizer. The reflection from the metasurface as well as the incident laser beam (to serve as a reference beam) are then directed to an infrared (IR) camera, creating interference fringe patterns. The incident laser beam is focused on the edge of the metasurface nanoantenna array. As a result, the scattered beam is reflected partly from the metasurface and partly from the Au backplane. This results in a lateral shift in the interference fringe patterns of the metasurface and the backplane when the applied bias is changed. We then fit these two cross-sections to sinusoidal functions and obtain the relative delay between the fitted sinusoidal curves when changing the applied voltage. The phase shift acquired due to the applied bias is then retrieved.[10] In the next step, to measure the reflectance the surface of the metasurface sample is illuminated by the NIR laser beam. Then, the beam reflected from the metasurface is guided to a spectrometer and the reflectance is calculated as

$$\text{Reflectance [\%]} = 100 \times \frac{R_{\text{MTS}} - R_{\text{dark}}}{R_{\text{reference}} - R_{\text{dark}}} \qquad (S5)$$

where R_{MTS}, $R_{\text{reference}}$, and R_{dark} are the raw reflectance from the metasurface sample, a mirror and the background, respectively
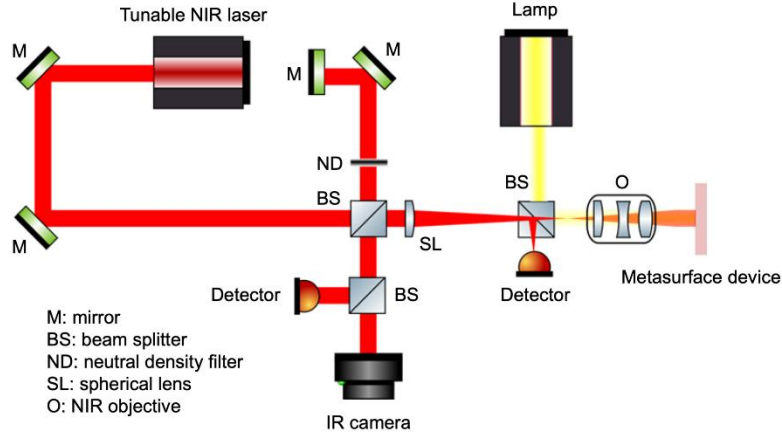
17

Figure S15. Optical setup used for the amplitude, phase, and beam steering measurements. The metasurface sample is illuminated by a tunable NIR laser. The reflected beam from the metasurface device is directed to a detector (amplitude measurement) and an IR camera (phase and beam steering measurements). The incident beam is also guided to the IR camera to be used as a reference for generation of the interference fringe patterns (phase measurement).

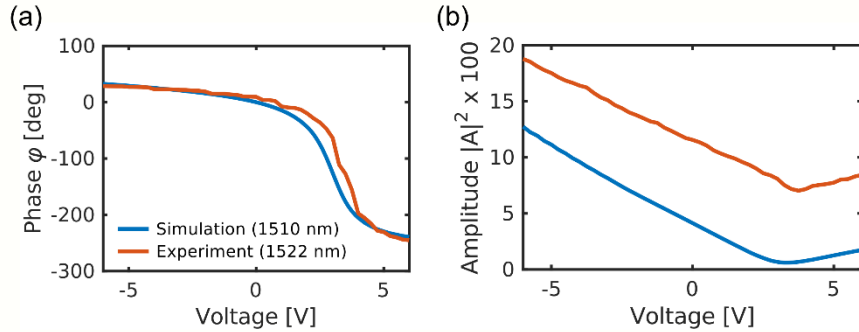## 14. Analytical model to account for experimental artefacts

Figure S16. (a) Phase shift and (b) reflectance for the electro-optically tunable metasurface introduced in Ref. 2 obtained with full-wave simulations ($\lambda = 1510$ nm) and experiments ($\lambda = 1522$ nm). The operational wavelength was shifted in experiments due to a change of the structural metasurface unit cell parameters post-fabrication.

Here, we discuss the changes that are made to the analytical model to reproduce experimentally measured beam steering radiation patterns. The data is based on simulations and experiments performed by Kafaie Shirmanesh *et al.*[2] Figure S16 shows a comparison of the simulated and experimentally measured phase and reflectance data. While the measured phase shift closely matches the simulated response, the experimentally measured reflectance R_{meas} is increased by an offset of approximately 7%. As discussed in the manuscript, this increase is attributed to a misalignment between the incident light polarization and the antenna, leading to enhanced specular reflection. In addition, the misaligned component of the incident light does not contribute to the phase accumulation and hence cannot be considered for optimization of the beam directivity. Therefore, we model the actual reflectance of the metasurface R_{actual} as R_{meas} - Δ_r with

18

Δ_r being a constant value that is determined as an average difference in reflectance over the applied bias range. To account for this change the intensity at $0°$ is increased by Δ_r.

Using the approximated reflectance of the metasurface, we computed the far-field radiation patterns for forward-designed four-level stairstep phase profiles. Figure S17 shows a comparison of the analytically predicted far-field radiation pattern to the experimentally measured beam steering performance for repetition numbers varying from RN = 1 (Fig. S17a) to 6 (Fig. S17f). We would like to note that in addition to the altered reflectance values, we also consider the change in the characteristic pitch size post-fabrication. By doing so, we are able to obtain an excellent match between the analytically predicted and experimentally measured beam steering performance. Small discrepancies in the sidelobe intensity are attributed to the fact that the adapted model is purely based on an approximate reflectance response of the metasurface. We further remark that due to the limited detectable angular range of our experimental setup, beam steering measurements could not be performed for RN = 1.
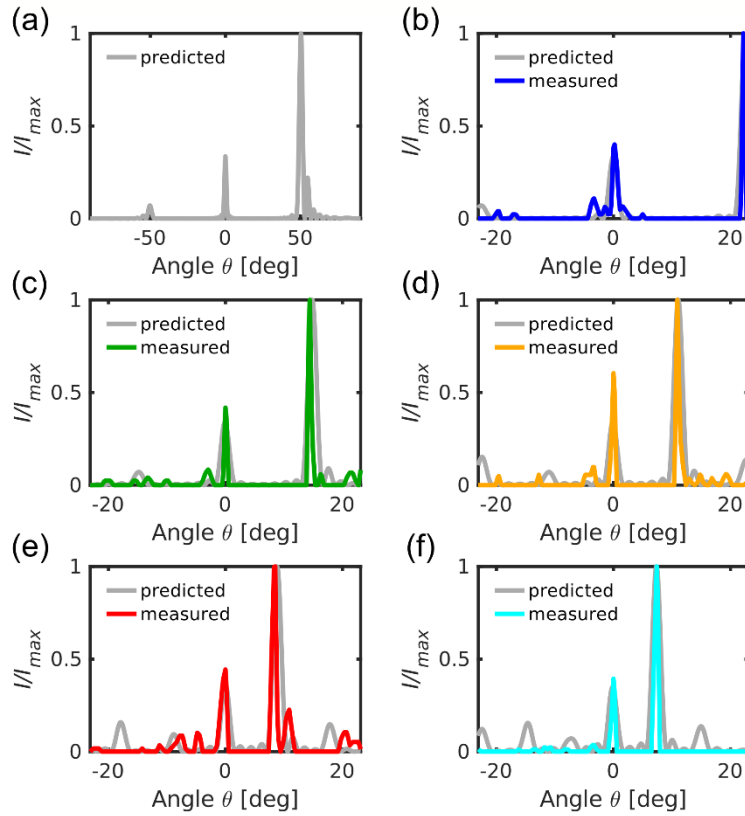

Figure S17. Analytically computed (grey) and experimentally measured (colored) far-field radiation patterns for forward-designed four-level stairstep phase profiles. Intensity I is normalized by its peak value I_{max}. The repetition numbers are varied as follows: (a) RN = 1, (b) RN = 2, (c) RN = 3, (d) RN = 4, (e) RN =5, (f) RN = 1. The operational wavelength is $\lambda = 1522$ nm and the characteristic pitch size of the experimental metasurface is continuously varying from 490 – 510 nm with the largest pitch size being at the center of the metasurface. Due to the limited detectable angular range of our experimental setup, beam steering measurements could not be performed for RN = 1.

15. Forward- and inverse-designed array profiles for experimental metasurface

Figure S18 illustrates the forward- and inverse-designed array profiles obtained using the optimization approach outlined in the manuscript. The analytically computed and experimentally measured far-field radiation patterns are shown in Fig. 6d and e, respectively.
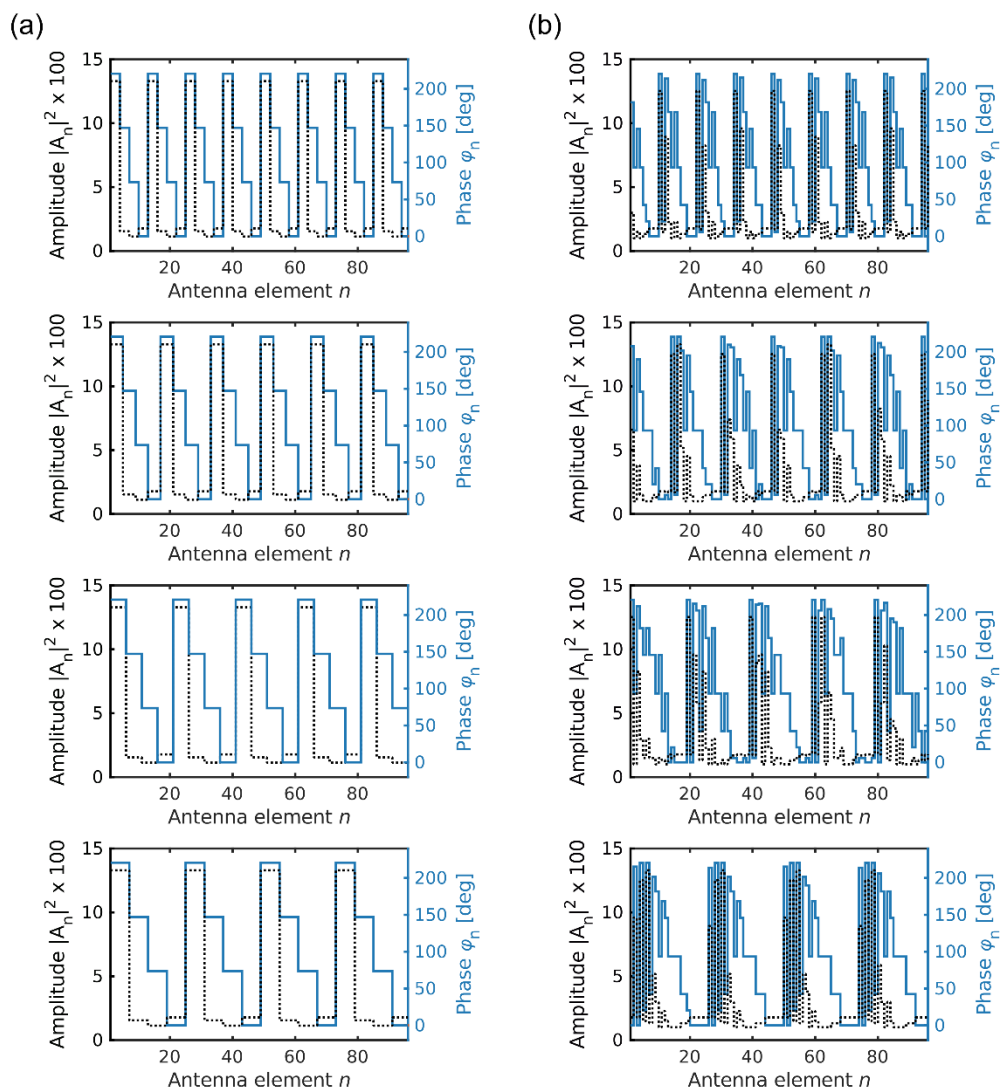

Figure S18. (a) Forward- and (b) respective inverse-designed spatial array amplitude (black) and phase (blue) profiles for the experimentally demonstrated beam steering results demonstrated in the manuscript. In the forward design case, the repetition number of the four-level stairstep phase profile is varied from RN = 3 (upper left) to 6 (bottom left).

16. Full-wave simulation of forward and inverse design for experimental metasurface array

In this section, we illustrate the far-field radiation patterns that are obtained using the simulated functional response of the experimental metasurface design presented in the manuscript (Fig. S14). In order to prevent breakdown of the gate dielectrics, the applied bias voltages were limited from -3.5 V to +3.5 V. The forward- and inverse-designed array phase and amplitude profiles are shown in Fig. S20a and b, respectively. The normalized intensity as a function of polar angle θ is computed using the analytical array factor equation as well as through full-wave simulations. It should be noted that due to the large size of the simulation region, a mesh size of 0.05 nm is used in the accumulation layer of the ITO film. The metasurface is illuminated by a plane-wave and the near-to-far field transformation is used to calculate the far-field radiation pattern of the beam steering metasurface. Fig. S20c and d indicate that for both the forward- and inverse-designed array profiles the analytically calculated radiation patterns match up with their simulated counterparts. Lower magnitude of the sidelobes at oblique angles is attributed to a non-unity antenna factor that is not accounted for in our calculations (see Supporting Information Part 1 for more details). The results of this study support the fundamental assumption of an independent scatterer model that is made when working with the array factor calculation.
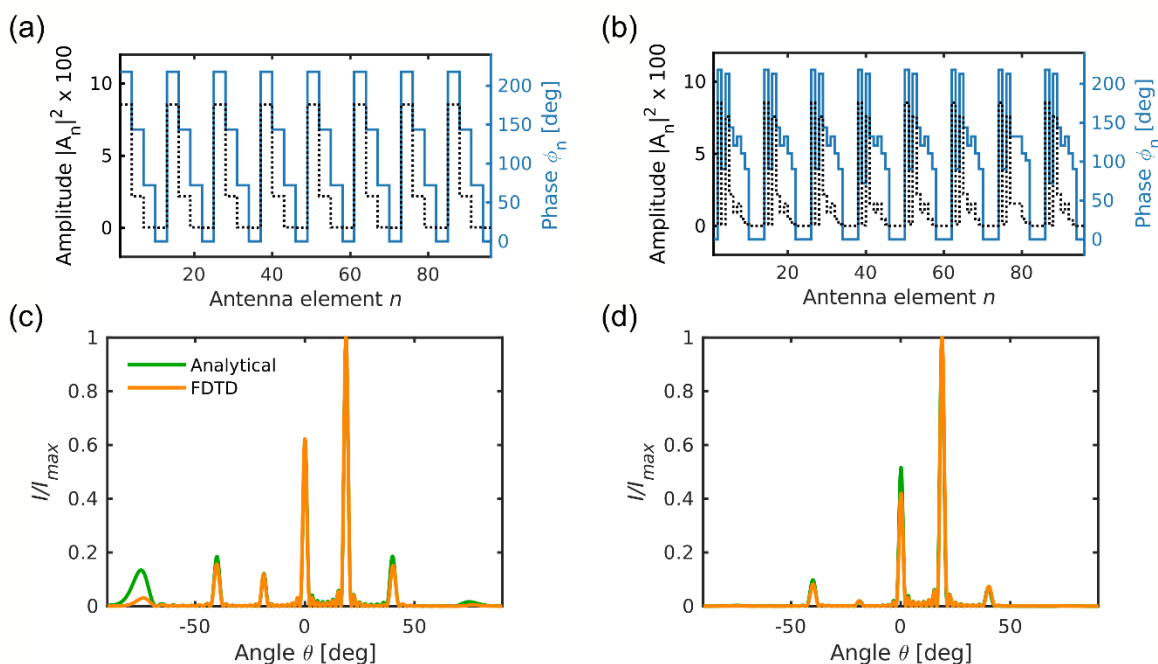

Figure S20. Array phase and amplitude profiles obtained using (a) a four-level stairstep forward design principle with a repetition number of RN = 3 and (b) an array-level inverse design approach. The corresponding far-field radiation patterns using analytical array factor calculations (green) and FDTD full-wave simulations (orange) are shown in (c) and (d), respectively. Results are illustrated for the experimental metasurface design presented in the manuscript. The operating wavelength is λ = 1545 nm.

17. Change in phase modulation with consecutive measurements

Figures S19a and b show the measured phase shift values obtained for two different metasurface samples at the wavelength of 1548 nm. The phase shift was obtained in three consecutive measurements performed on each sample. As can be seen in Fig. S19a, in the first round of the phase measurement, a phase shift of 201° was obtained for the first test sample. In the second and the third rounds of the phase measurement, the phase shift provided by the first sample decreased to 193° and 187°, respectively. As a result, a reduction of 4% and 3.1% in the phase modulation, respectively, was observed when the phase measurement was repeated each time. Figure S19b also shows that for phase measurements performed on the second sample, phase shifts of 196°, 182°, and 171° were obtained in the first, second, and third round. This amounts to a decrement of 7.1% and 6% in the phase shifts attained when repeating the measurement. As can be seen in Fig. S19, applying the DC bias to the metasurface samples led to a reduction of the phase modulation provided by the metasurface. This can be attributed to a degradation of the gate dielectric under an applied bias as a result of the existence of pin-holes. It should also be noted that the extremum applied biases were chosen to be ±3.9 V, and ±4.2 V for the first and second metasurface samples, respectively, in order to prevent the gate dielectrics from breakdown.
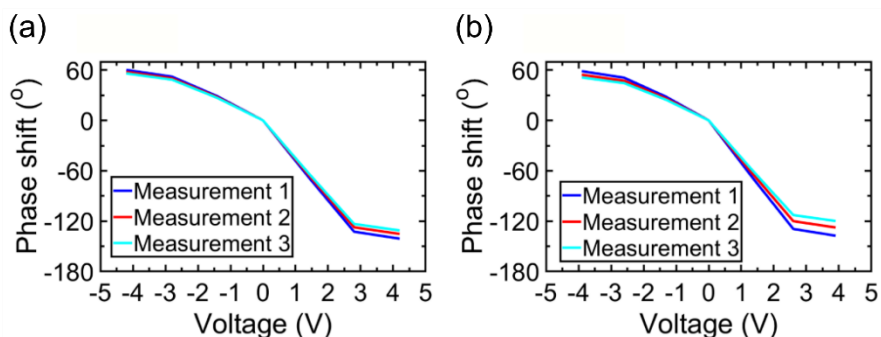

Figure S19. Phase shift measured on (a) the first and (b) the second metasurface test sample under repetitive phase measurements. A maximum phase shift of 201°, 193°, and 187° was obtained for the first test sample in the first, second, and the thirds round of the phase measurement, respectively. The phase modulation provided by the second test sample was measured to be 196°, 182°, and 171° in three consecutive measurements.

18. Error tolerance of forward designs in non-ideal antenna arrays

Figure S21 shows the phase error tolerance for forward- and inverse-designed arrays steering at $\theta_r = 18.3°$. To obtain the threshold phase error tolerance, we consider the relative change in directivity for error introduced into the entire array design ($f = 100\%$). The threshold performance of $0.9 \times D_{\delta=0°}$ is obtained for larger amounts of phase disorder δ in forward designs: $\delta_{inverse} = 60°$ $< \delta_{forward,lin} = 100° < \delta_{forward,step} = 140°$. Here, $D_{\delta=0°}$ is the beam directivity with the respective array design without any introduction of phase noise. The increased error tolerances in forward designs are justified by the uniform phase shifts between adjacent antennas. As a consequence, forward designs tolerate larger errors before reaching substantial performance losses. It is to be noted that directivities of stairstep profiles can surpass $D_{\delta=0°}$ for $\delta \leq 40°$. Since stairstep designs represent

simplified gradient phase profiles, small amounts of phase disorder can lead to closer resemblances to higher-directivity linear array designs.
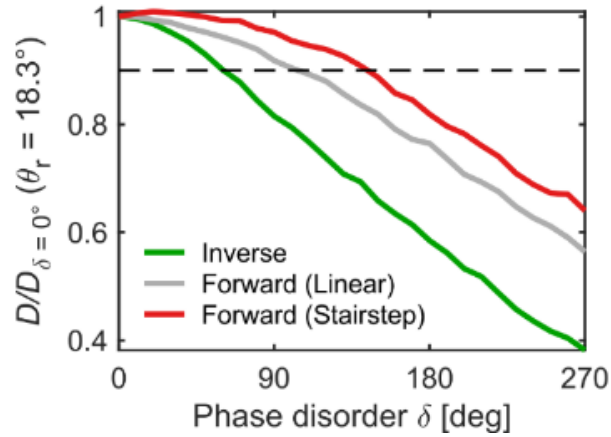

Figure S21. Relative change in directivity $D/D_{\delta=0°}$ for increasing phase disorder δ for steering at θ_r = 18.3°. $D_{\delta=0°}$ is the directivity of the respective array design prior to any introduction of phase noise. Results are illustrated for inverse-designed arrays (green), as well as forward-designed linear (grey) and stairstep (red) phase profiles for the non-ideal active metasurface.[2] The black dashed line marks the threshold directivity at which the disordered beam directivity drops to $0.9 \times D_{\delta=0°}$. The change in directivity is obtained as an average over 100 implementations.

REFERENCES

1. Balanis, C. A. *Antenna Theory: Analysis and Design*; Wiley, **2016**.

2. Kafaie Shirmanesh, G.; Sokhoyan, R.; Wu, P. C.; Atwater, H. A. Electro-Optically Tunable Multifunctional Metasurfaces. *ACS Nano* **2020**, 14, 6912–6920.

3. Wang, J.; Hou, P.; Cai, H.; Sun, J.; Wang, S.; Wang, L.; Yang, F. Continuous Angle Steering of an Optically- Controlled Phased Array Antenna Based on Differential True Time Delay Constituted by Micro-Optical Components. *Opt. Express* **2015**, 23, 9432-9439.

4. Engström, D.; Bengtsson, J.; Eriksson, E.; Goksör, M. Improved Beam Steering Accuracy of a Single Beam with a 1D Phase-Only Spatial Light Modulator. *Opt. Express* **2008**, 16, 18275-18287.

5. Mishra, S.; Yadav, R. N.; Singh, R. P. Directivity Estimations for Short Dipole Antenna Arrays Using Radial Basis Function Neural Networks. *IEEE Antennas Wirel. Propag. Lett.* **2015**, 14, 1219–1222

6. Saleem, S. S.; Ahmed, M. M.; Rafique, U.; Ahmed, U. F. Optimization of Linear Antenna Array for Low SLL and High Directivity. *Proc. 2016 19th Int. Multi-Topic Conf. INMIC, IEEE* **2016**, 1–6.

7. Li, S. Q.; Xu, X.; Veetil, R. M.; Valuckas, V.; Paniagua-Domínguez, R.; Kuznetsov, A. I. Phase-Only Transmissive Spatial Light Modulator Based on Tunable Dielectric Metasurface. *Science* **2019**, 364, 1087–1090.

8. Komar, A.; Fang, Z.; Bohn, J.; Sautter, J.; Decker, M.; Miroshnichenko, A.; Pertsch, T.; Brener, I.; Kivshar, Y. S.; Staude, I.; Neshev, D. N. Electrically Tunable All-Dielectric Optical Metasurfaces Based on Liquid Crystals. *Appl. Phys. Lett.* **2017**, 110, 071109.

9. Wu, P. C.; Pala, R. A.; Kafaie Shirmanesh, G.; Cheng, W.; Sokhoyan, R.; Grajower, M.; Alam, M. Z.; Lee, D.; Atwater, H. A. Dynamic Beam Steering with All-Dielectric Electro-Optic III-V Multiple-Quantum-Well Metasurfaces. *Nat. Commun.* **2019**, 10, 1–9.

10. Kafaie Shirmanesh, G.; Sokhoyan, R.; Pala, R. A.; Atwater, H. A. Dual-Gated Active Metasurface at 1550 nm with Wide (>300°) Phase Tunability. *Nano Lett.* **2018**, 18, 2957–2963.

11. Forouzmand, A.; Salary, M. M.; Kafaie Shirmanesh, G.; Sokhoyan, R.; Atwater, H. A.; Mosallaei, H. Tunable All-Dielectric Metasurface for Phase Modulation of the Reflected and Transmitted Light *via* Permittivity Tuning of Indium Tin Oxide. *Nanophotonics* **2019**, 8, 415–427.

12. Sokhoyan, R.; Grajower, M.; Feigenbaum, E.; Elhadj, S.; Atwater, H. A. Electrically Tunable Conducting Oxide Metasurfaces for High-Power Applications. *In preparation.* **2020**.

13. Johnson, J. M.; Rahmat-Samii, Y. Genetic Algorithms in Engineering Electromagnetics. *IEEE Antennas Propag. Mag.* **1997**, 39, 7–21.