Appendix A

ARRAY-LEVEL INVERSE DESIGN

A.1 Iterative genetic optimization

The following pages provide further details on the implementation and robustness of the results obtained with the iterative genetic optimization approach introduced in Chapter 2. We would like to note here that a variety of global optimization algorithms, including particle swarm optimization, simulated annealing, and others can be used for the purpose of array-level optimization.

Numerical framework

Figure A.1 shows a flowchart of the optimization approach implemented in our study using the global optimization toolbox on MATLAB R2018b. For beam steering, 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. Here, 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. 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



Figure A.1: Flowchart of iterative genetic optimization for an array of N_{tot} = 96 antennas. The inner loop represents the iterative genetic optimization with increasing number of optimization variables to approach the high dimensionality of the underlying optimization problem. The outer loop describes a series of optimization rounds that allow to take the optimal solution over multiple repetitions.

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 Eq. (2.3).

Optimization parameters

A matrix containing the metasurface-specific tunable phase-amplitude relation is provided as an input to the algorithm. Then, a discrete integer optimization is performed including the matrix rows (that define the scattered light response for various voltages). The lower and upper bounds are variable and define the first and last row of the matrix, respectively. The default creation, crossover, and mutation functions of discrete genetic optimization enforce each variable to be an integer, as discussed in [338]. To accommodate for the large number of parameters that are to be optimized, the population size is increased to 200. The optimized results are obtained with a crossover fraction of 0.95 and an elite count of 20. Each iteration stops when the average change in the best function value over 250 generations is less than 10^{-6} .

Computational cost

In contrast to forward-designed array profiles that rely on an analytical equation (Eq. 2.3), the inverse design approach comes with enhanced computational cost due to a consideration of the antenna-specific functional response. For the problems analyzed in this work, the optimal solution in each iteration is reached 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 A.2 shows the average computation time with 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 angles $(\theta_r = 9.0^\circ, 10.9^\circ, 13.6^\circ, 18.3^\circ, 28.1^\circ, 70.7^\circ)$. Notably, the computation time scales linearly as $O(N_{tot})$ in the investigated regimes while the solution space scales exponentially as $O(s^{N_{\text{tot}}})$ 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 through full-wave simulations [42]. 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. In future work, the stopping criteria can be optimized such that the computational cost is reduced without a significant loss in best performance.

Robustness of optimization algorithm

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



Figure A.2: Averge computation time TOC_{avg} as function of optimization variables N_{tot} . 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.

(crossover and mutation) further introduce variability into population members. Due to the stochastic operations in genetic optimization, convergence characteristics differ between individual optimization rounds [71]. 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 A.3a shows the range of optimized directivities for the three analyzed methods for a steering angle of $\theta_r = 18.3^\circ$. 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 forwarddesigned linear phase profiles results in a maximal increase in directivity of 77% compared to the previously demonstrated stairstep forward design with $D_{\text{forward}} =$ 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 Section 2.3. 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



Figure A.3: **Robustness of optimization algorithm.** (a) Distribution of the optimized directivity over $r_{tot} = 20$ optimization rounds for a steering angle of $\theta_r = 18.3^{\circ}$. The results are illustrated for three different optimization methods: a direct, noniterative 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.

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. A.3b). 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. A.3c) leads to stronger directivity enhancements due to an unbiased and thus more extensive exploration space. Finally, the iterative approach 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. A.3a. For the case studied here, 75% of the optimized performances lie within 3% of the maximal directivity $D_{opt,iter} = 72.7$ (Fig. A.3d).

A.2 Metasurface fabrication

To fabricate the metasurfaces used in Chapter 2 of this thesis, we followed the fabrication procedure for the ITO-based field-effect tunable metasurface by Kafaie Shirmanesh *et al.* [42]. First, we clean silicon (Si) substrates with a 1 μ m thick silica (SiO_2) layer on top using standard cleaning processes. Then the outermost parts of the connection pads are patterned using photolithography. After developing the exposed photoresist, we deposit a 20 nm thick titanium (Ti) layer followed by a 200 nm thick gold (Au) layer via e-beam evaporation. The excess resist and the Ti/Au film are removed through a lift-off process in acetone. Then the Au back reflector is patterned using electron-beam lithography (EBL) [VISTEC EBPG 5000+] at an acceleration voltage of 100 keV after spinning an e-beam resist layer. The exposed e-beam resist layer is then developed and 3 nm of chromium (Cr) followed by an 80 nm-thick Au layer are deposited using an e-beam evaporator. In a next step, a 9.5 nm-thick alumina (Al_2O_3) layer is deposited on the samples through shadow mask via EBL. Once the exposed resist is developed, we sputter a 5 nm thick ITO layer via room-temperature RF magnetron sputtering. The deposition pressure is 3 mTorr and the applied RF power is 48 W. The plasma is struck by using argon (Ar) gas with a flow rate of 20 sccm, and argon/oxygen gas (Ar/O₂:90/10) with a tunable flow rate is used to control the carrier concentration of the deposited layer [53]. After lifting off the excess e-beam resist and films, the contact pads of the ITO layer are patterned via EBL followed by a deposition of a Ti/Au film (20 nm/200 nm) using an e-beam evaporator. We then deposit the hafnium oxide/aluminum oxide laminate (HAOL) gate dielectric layer through shadow masks by ALD [53]. In the next step, the nanoantennas as well as the inner contact lines are patterned on the sample using EBL. After developing the exposed e-beam resist, a 2 nm-thick germanium (Ge) layer followed by a 40 nm thick Au layer is deposited on the sample using an e-beam evaporator. Once the lift-off process is done, a 60 nm thick SiO_2 layer is deposited as the top coat through shadow mask via e-beam evaporation. Finally, 96 metasurface element pads and 4 ITO pads are wire-bonded from the sample to 100 conducting pads on a sample mounting printed circuit board (PCB) which itself is controlled by using a voltage-driving PCB [42].

A.3 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 A.4 shows the optical setup used to



Figure A.4: **Optical setup for amplitude, phase and beam steering measurements.** The metasurface sample is illuminated by a tunable NIR laser. The reflected beam from the metasurface 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.

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 is illuminated by a tunable NIR laser which is focused on the sample by an objective with a long working distance (Mitutoyo M Plan Apo 20x, 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 [53]. In the next step, to measure the reflectance, the surface of the metasurface 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



Figure A.5: Comparison of simulated and experimentally measured optical response of ITO metasurface. (a) Phase shift and (b) reflectance for field-effect tunable metasurface introduced in [42] obtained through full-wave simulations ($\lambda = 1510 \text{ nm}$) and experiments ($\lambda = 1522 \text{ nm}$).

Reflectance [%] =
$$100 \times \frac{R_{\text{MTS}} - R_{\text{dark}}}{R_{\text{reference}} - R_{\text{dark}}}$$
 (A.1)

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

A.4 Analytical model accounting for experimental artifacts

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.* [42]. Figure A.5 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%. 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 $R_{\text{meas}} - \Delta_{\text{r}}$ with 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 A.6 shows a comparison of the analytically predicted far-field radiation pattern to the



Figure A.6: Predicted and measured far-field radiation patterns with adapted model. Analytically computed (gray) and experimentally measured (colored) normalized intensity I/I_{max} vs polar angle θ for forward-designed stairstep phase profiles with repetition numbers of (a) RN = 1, (b) RN = 2, (c) RN = 3, (d) RN = 4, (e), RN = 5, (f) RN = 6. The operational wavelength is $\lambda = 1522nm$.

experimentally measured beam steering performance for repetition numbers varying from RN = 1 to 6. We would like to note that in addition to the altered reflectance values, we also consider a continuously varying change in the characteristic pitch size of the metasurface from 490 nm – 510 nm, with the largest pitch size being at the center of the metasurface. 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 in our experimental setup, beam steering measurements could not be performed for RN = 1.

Appendix B

SPACE-TIME MODULATED METASURFACES

B.1 Active metasurface phase measurements

Figure B.1 plots the maximum phase shift achieved by our active metasurface as a function of wavelength. The largest phase shift occurs between 1465 nm and 1470 nm, corresponding to the minimum reflectance measured in Fig. 3.3d. The phase shift of this device at the design wavelength for this work (1530 nm) is near-zero ($< 20^{\circ}$). While there is larger amplitude and phase modulation at shorter wavelengths (~ 1470 nm), we chose to work at 1530 nm because it is the shortest wavelength we could access with sufficient power in our experimental setup. This work could be repeated closer to resonance using the same principles discussed here and the final results would be identical, but the overall efficiency of the device would be higher. Specifically, there would be a higher ratio of total modulated light to unmodulated light. However, since our measurement technique already filters out unmodulated light, in this work we are still able to show the same principles of space-time modulation and engineer the power sent to each modulated frequency with good signal-to-noise ratio (SNR).

It is noteworthy that this work demonstrates space-time diffraction with excellent directivity at each modulated frequency using a device with such limited reflectance and phase modulation. If this device was modulated in the quasi-static regime, it would exhibit very low diffraction efficiency. By using space-time modulation, roughly the same total amount of light is diffracted as in the quasi-static case, but the light sent to the +/- 1st orders are separated in frequency from the 0th order (normally reflected light) which allows for easier detection with high SNR. This is a major benefit of space-time metasurfaces.

B.2 Waveform optimization based on quasi-static metasurface response

In Section 3.3, we discussed how a real-time optimized waveform can be used to generate desired frequency spectra. An alternative approach would be to use the quasi-static metasurface response and perform an analytical optimization to design tailored waveforms. This method follows a similar principal to the optimization techniques introduced in Chapter 2, but instead of optimizing in the spatial domain, we optimize the covarying amplitude and phase in time.



Figure B.1: Quasi-static phase shift of two-electrode active metasurface. The plotted phase shift is the absolute value of the change in phase between an applied bias of -6 V and 6 V.

To formulate the waveform, we express it as a polynomial function $V(t) = \sum_{n} a_{n} \cdot t^{n-1}$ where n - 1 represents the degree of the polynomial. The optimization minimizes a figure of merit (FOM) defined as $2 \cdot \sum_{i} p_{i}/p_{1} + \text{mean}(V(t))$, where p_{i} is the power at the *i*-th harmonic. Initial optimizations with varying polynomial degrees indicate that a 7-th degree polynomial results in the highest ratio between the $+1^{\text{st}}$ and 0^{th} harmonic while maintaining minimal mean voltage. It is desirable to maintain a mean voltage close to 0 V to avoid accumulation of charge carriers in the ITO layer over time. We additionally impose a constraint to have equal values at the beginning and end of a waveform period, preventing sudden voltage jumps that would be challenging to realize at high frequencies.

For illustration, we performed an optimization for a metasurface operating at 1555 nm using its quasi-static reflectance and phase response (Fig. B.2a). The resulting optimized waveform is given by

$$V(t) = 1.6 + 1.4 \cdot t - 8.4 \cdot t^2 - 2.3 \cdot t^3 + 3.4 \cdot t^4 - 9.9 \cdot t^5 + 6.4 \cdot t^6 + 9.4 \cdot t^7.$$
 (B.1)

The mean voltage of this waveform (Fig. B.2b) is 0.1176 V, while the suppression ratio of p_1/p_0 is 19.96 (Fig. B.2c).

Despite the high suppression ratio predicted in calculations, experimental results using a 10 kHz waveform did not exhibit suppression of the unmodulated, specularly



Figure B.2: **Polynomial waveform optimization based on quasi-static response.** (a) Experimentally measured quasi-static reflectance and phase response (points) and fits to experimental data (dashed) used for optimization. (b) Optimized polynomial voltage waveform as function of time. The time-axis is labeled in arbitrary units as the 100 points can be spaced apart at desired increments to obtain target frequencies. (c) Conversion efficiency as function of frequency for optimized waveform at 10 Hz.

reflected light. This discrepancy suggests that the reflectance and phase response of the metasurface in the quasi-static regime differ from those attainable at high frequencies. While high-frequency reflectance can be measured using an oscilloscope, alternative techniques are required to accurately characterize the phase response at target modulation frequencies. To address this limitation, we integrated a real-time optimization algorithm in our experimental setup, as shown in Fig. 3.4. Although real-time optimization requires longer processing times due to iterative measurements, it enables significantly improved performance for a variety of target frequency spectra [95].

Appendix C

BARIUM TITANATE-BASED ACTIVE METASURFACES

C.1 XRD scans for bulk substrates and spalled film

Commercially bought BTO substrates [339] were produced via top-seeded solution growth and were single-crystalline with multiple domains. We confirmed the singlecrystallinity of bulk substrates and spalled films through $\theta - 2\theta$ x-ray diffraction (XRD) scans. As shown in Fig. C.1, the prominent peaks for both the bulk substrate and spalled thin film correspond to the (001) and (100) orientations, along with their higher-order diffractions. These peaks arise due to the multi-domain nature of the substrates and films, combined with the millimeter-scale x-ray probing spot. To probe the domain structure at a finer scale, we employed electron backscatter diffraction (EBSD), as discussed in Fig. 4.3. Similar results were obtained for (100)-oriented crystals.

Figure C.1b shows the $\theta - 2\theta$ scan for a thin film spalled from a (001)-oriented singlecrystal. The reduced intensity observed for the thin film is primarily attributed to its surface roughness and reduced thickness. In addition to the signal from BTO, we also measure a minimal contribution from the underlying Ni layer due to the size and position of the x-ray probing spot relative to the spalled film area.

C.2 Reusability of single-crystal BTO substrates through repolishing

One major advantage of spalling compared to alternative growth techniques is the ability to reuse bulk substrates by introducing a polishing step between spalls. Considering that the substrates are typically 0.5 - 1 mm thick, and each iteration removes less than 50 μ m of material thickness, the effective cost of a spall can be lowered up to 20-fold, bringing the expense down to tens of dollars per square centimeter.

In this study, various polishing procedures were tested and the RMS roughness r of each was compared to that of as-bought substrates. Figure C.2a shows an AFM scan of an as-bought substrate with r = 5.9 nm. Hand polishing of BTO substrates yields planarized surfaces but results in a significantly higher RMS roughness of r = 66.4 nm. While spalls can still be performed with this level of roughness, it is insufficient for many applications.



Figure C.1: **XRD data for (001)-oriented bulk substrate and spalled thin film.** $\theta - 2\theta$ scan for (a) bare bulk substrate and (b) spalled thin film on a Ti/Au seed layer and electroplated Ni. BTO diffraction orders corresponding to various crystal orientations are labeled in both figures, Ni diffraction orders are indicated in gray in (b).

To address this, we explored vibration polishing using a slurry of silica (SiO₂) or alumina (Al₂O₃) nanoparticles, which significantly reduce surface roughness. Figures C.2b and c show AFM scans of the BTO substrate after vibration polishing with 60 nm-sized SiO₂ particles for 7 hours and 50 nm-sized Al₂O₃ particles for 3 hours, respectively. The resulting RMS roughness values were r = 27.2 nm and r = 9.4 nm. The stronger reduction in surface roughness achieved with Al₂O₃ particles is attributed to their higher material hardness relative to BTO.

It is important to note that while vibration polishing does not planarize the substrate, it effectively restores surface roughness to near-initial levels, enabling large-scale spalls suitable for device integration, as shown in Fig. 4.1b.

C.3 Surface quality of (100) and (001)-oriented bulk substrates

Commercially bought single crystal BTO substrates [339] are produced using topseeded solution growth (TSSG). This process involves growing boules from molten solutions at elevated temperatures. During this process, the controlled cooling and crystallization of the molten material enable the formation of high-quality



Figure C.2: **AFM scans of BTO substrate as-bought vs polished after spalling.** (a) As-bought BTO substrate, (b) as-bought BTO substrate, vibration polished with 60 nm-sized SiO₂ particles for 7 hours, (c) BTO substrate post-spall, vibration polished with 50 nm-sized Al₂O₃ particles for 3 hours. Scale bars: 10 μ m.

single crystals [340]. After the boule is formed, sections of it are cut and polished to create substrates. These substrates are typically multidomain in nature, with distinct regions corresponding to different orientations of the polar axis (spontaneous polarization).

This multidomain structure forms as BTO undergoes a structural phase transition from the cubic to tetragonal phase upon cooling below its Curie temperature. This transition generates spontaneous polarization, with the formation of domains being influenced by various factors such as the cooling rate, internal and external stresses, and thermal gradients during TSSG [341]. These conditions lead to the formation of 90° and 180° domain walls, which separate regions with differing polarization orientations. Further details on the domain structure of BTO can be found in Refs. [341–345].

We note significant differences in the bulk substrate quality between (100)- and (001)-oriented crystals, as illustrated in Fig. C.3. The (100)-oriented substrates feature prominent 90° domain patterns and finer surface ridges (domain bundles). In contrast, (001)-oriented substrates exhibit a more uniform appearance with fewer domain boundaries. This discrepancy likely arises from the anisotropic strain in the in-plane directions of *a*-axis crystals during cooling, which promotes the formation of additional domain walls. The higher domain density in (100)-oriented substrates results in increased surface roughness. The elevated roughness carries over to spalled thin films, leading to a more irregular topography. By contrast, the smoother surface of (001)-oriented substrates yields spalled films with more controllable and uniform surface properties.



Figure C.3: **Single crystal BTO substrates with domains.** Optical microscope image of (a) (100)-oriented single crystal substrate with domains, and (b) (001)-oriented single crystal substrate with domains. Scale bar: 2.5 mm.