



Research article

Dong Cheon Kim*, Andreas Hermerschmidt, Pavel Dyachenko and Toralf Scharf

Inverse design of diffractive optical elements using step-transition perturbation approach

<https://doi.org/10.1515/aot-2020-0046>

Received July 31, 2020; accepted September 17, 2020;
published online October 7, 2020

Abstract: Diffractive optical elements are ultra-thin optical components required for a variety of applications because of their high design flexibility. We introduce a gradient-based optimization method based on a step-transition perturbation approach which is an efficient approximation method using local field perturbations due to sharp surface profile transitions. Step-transition perturbation approach be available to calculate the gradient of figure of merit straightforwardly, we implemented optimization method based on this gradient. This fast and accurate inverse design creates binary (2-level) diffractive elements with small features generating the wide angle beam arrays. The results of the experimental characterization confirm that the optimization based on the perturbation method is valid for 1-to-117 fan-out grating generating beam pattern of linear array.

Keywords: diffractive optical elements; fan-out gratings; gradient-based optimization; inverse design; step-transition perturbation approach.

1 Introduction

Diffractive optical elements (DOEs) are used for a variety of optical systems because of their compact size, high design flexibility, and ease of mass production. A good example of this type of device is fan-out grating, often also referred to as diffractive beam-splitter which creates multi spots by deflecting an incident light into different diffraction orders.

*Corresponding author: **Dong Cheon Kim**, Nanophotonics and Metrology Laboratory (NAM), EPFL, 1015 Lausanne, Switzerland, E-mail: dong.kim@epfl.ch. <https://orcid.org/0000-0002-4234-3840>
Andreas Hermerschmidt and Pavel Dyachenko, Holoeye Photonics AG, 12489 Berlin, Germany, E-mail: andreas.hermerschmidt@holoeye.com (P. Dyachenko)
Toralf Scharf, Nanophotonics and Metrology Laboratory (NAM), EPFL, 1015 Lausanne, Switzerland, E-mail: toralf.scharf@epfl.ch

Applications of fan-out elements include multi-focal microscopy [1, 2], camera calibration [3], optical system distortion measurement [4], optical interconnects [5], and structured light projectors [6, 7]. In particular, wide angle DOEs are used in promising field of applications [7–9] with recent progress in fabrication technology realizing nano-scale features. When only small diffraction angles are required, the iterative Fourier transform algorithm (IFTA) [10–12] based on the thin element approximation (TEA) [13] is widely and successfully used for the design of the microstructure surface of the DOEs. However, this approach suffers from several severe shortcomings and is no more valid when larger angles are required [14]. One of the most severe problems is the insufficient modeling of the light field transmission through the DOE by the TEA [15]. A precise modeling can be obtained by rigorous electromagnetic diffraction theory, such as rigorous coupled-wave analysis (RCWA) [16–18]. However, the use of parametric optimization based on the rigorous analysis is often computationally heavy because the gratings with many parameters are leading to high-dimensional optimization problems. To overcome this limitation, we employ gradient-based algorithms [19] based on the step-transition perturbation approach (STPA) [20]. In gradient-based optimization of optical elements, within the overall design parameter space, one iteratively searches the maximum or minimum point based on calculated gradient [21, 22].

STPA enables rapid approximate calculation of diffraction intensity of wide angle DOEs, which is an approximate method based on local field perturbations generated by sharp transitions of the surface profile of diffractive elements. We can describe analytically the gradient of diffraction efficiency with respect to the design variables of optical elements using STPA. It thus is feasible to calculate the gradient straightforwardly with accuracy as much as the approach based on the rigorous electromagnetic diffraction theory if most of the features of the structure are bigger than the wavelength of the incident light.

Here, we introduce a design approach based on STPA for wide angle DOEs, yielding improvements in the uniformity of the created patterns while maintaining the total

diffraction efficiency. We have focused our efforts on designing binary (*i.e.*, 2-level) micro-structures because they are most easy to fabricate and thus obviously are very attractive for optical systems.

A schematic of DOE generating linear spot array is shown in Figure 1(a). In order to obtain a starting condition for our optimization, a one-dimensional (1D) diffractive fan-out DOE is designed by TEA-based IFTA (see Figure 1(b) inset). Figure 1(b) present the performance of this initial grating as a function of diffraction angle. Obviously, the performance of this initial fan-out element is very unsatisfactory with respect to uniformity even when the maximum diffraction angle is over 7° . We then apply our method to optimize large-angle 1D diffractive beam splitters but also compare to the results from parametric optimization based on RCWA. We measured the performance of fabricated samples based on the optimized designs and compared to calculated diffraction efficiency using RCWA simulation. Based on this analysis, we can show the excellent strengths of our design method.

2 Inverse design methods

An important aspect of the optimization process is the parametrization used to describe the shape of the optical elements, which can significantly affect the performance and computational cost. Figure 1 illustrates an example of a 1D binary phase grating profile with $2K$ transitions in position x_k within a single grating period. We use these positions of transition points as the set of design parameters $x = [x_1, \dots, x_k, \dots, x_{2k}]$ and define the figure of merit (FOM) to optimize DOEs creating diffraction pattern with uniform intensity distribution:

$$F(x) = \sum_{m=-M}^M [\eta_m(x) - \eta_{obj}]^2 \quad (1)$$

where F represents the difference between the calculated diffraction efficiency η_m and the target diffraction efficiency η_{obj} in diffraction orders. The gradient of the FOM with respect to transition positions $\nabla_x F$ is crucial in determining the search direction to optima. For example, if the total number of transitions $2K$ is large, it may easily become computationally heavy to calculate the gradient by RCWA analysis. The STPA, however, allows expressing the variation for a diffraction efficiency with respect to transition positions as an analytical solution so that it can calculate the gradient straightforwardly.

It has been reported by T. Vallius et al. [20] that in fact the approximated method based on local field

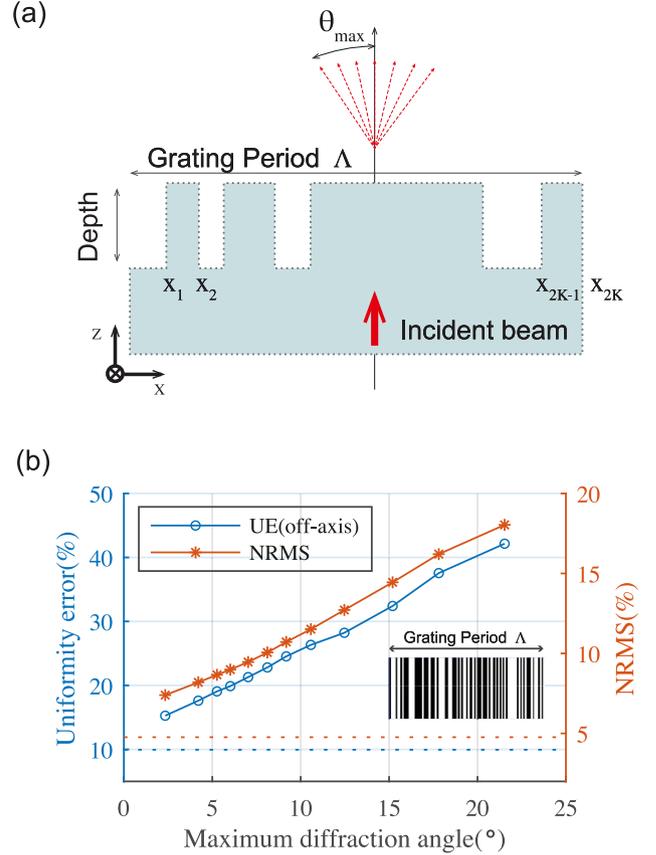


Figure 1: (a) The surface profile of a grating structure in a single period. (b) Uniformity of 1-to-117 diffractive beam splitter designed by TEA-based IFTA as a function of maximum diffraction angle when the grating period Δ decreases. The uniformity error and normalized root-mean-square error (NRMS) calculated using Eqs. (9) and (10). The dashed lines indicate the uniformity values calculated based on TEA. The insets show the layout of the single unit cell with total 60 transition points.

perturbations from sharp step-transitions enables rapid calculation of diffraction patterns of DOEs in the non-paraxial domain. The perturbations are observed in the field distribution directly after sharp vertical transitions of binary gratings. The TEA calculation, however, yields a constant amplitude and phase. This omission of perturbations in TEA makes computing inaccurate gratings with wavelength-scale structures, *i.e.*, the gratings creating the wide angle arrays. Thus, we can accurately calculate the diffraction efficiency using the model which combines the TEA with field disturbances caused by sharp transitions in the surface profile calculated by RCWA. We define the field perturbation behind the k th sharp transition located at the point x_k in the surface profile as

$$p_k(x) = \begin{cases} U_k^R(x) - U_k^T(x) & \text{if } |x| < \Delta_T \\ 0 & \text{elsewhere} \end{cases} \quad (2)$$

where $U_k^R(x)$ and $U_k^T(x)$ are field calculated by RCWA and TEA, respectively and Δ_T is the truncation parameter that is chosen 10λ in the calculations [20]. In Figure 2(a), (b), the amplitude and phase of the field distribution directly after an isolated step transition determined by TEA and RCWA is presented. The field perturbations of binary gratings consist of only two kinds of oscillation corresponding to left-side and right-side transition point in a ridge. Therefore, the constructed field behind binary grating with many transition points is described by the x -axis shifts of the two field perturbations $p_1(x)$ and $p_2(x)$ in the following expression:

$$\begin{aligned} U(x) &= U^T(x) + \sum_{k=1}^{2K} p_k(x) \\ &= U^T(x) + \sum_{k=1}^K p_1(x - x_{2k-1}) + \sum_{k=1}^K p_2(x - x_{2k}) \end{aligned} \quad (3)$$

where $2K$ is the total number of the transitions. The amplitude and phase of the field perturbation of right-side of a ridge $p_2(x)$ is represented in Figure 2(c), (d). The diffraction amplitude of m th order in far field is given by m th Fourier coefficient of $U(x)$ as

$$\begin{aligned} A_m &= \frac{1}{\Lambda} \int_0^\Lambda U(x) \exp(-i2\pi mx/\Lambda) dx \\ &= T_m + D_m \end{aligned} \quad (4)$$

where Λ is the grating period and T_m and D_m is the Fourier coefficient of the field calculated by TEA and a field perturbation contribution, respectively.

$$T_m = \frac{1}{\Lambda} \int_0^\Lambda U^T(x) \exp(-i2\pi mx/\Lambda) dx \quad (5)$$

$$\begin{aligned} D_m &= P_m \sum_{k=1}^K \exp(-i2\pi mx_{2k-1}/\Lambda) \\ &\quad + P_{-m} \sum_{k=1}^K \exp(-i2\pi mx_{2k}/\Lambda) \end{aligned} \quad (6)$$

where the Fourier coefficient P_m of field perturbation $p_1(x)$ is expressed as

$$P_m = \frac{1}{\Lambda} \int_0^\Lambda p_1(x) \exp(-i2\pi mx/\Lambda) dx \quad (7)$$

The Fourier coefficient of $p_2(x)$ is P_{-m} in Eq. (6) because the $p_2(x)$ is an even function of $p_1(x)$. We found D_m in Eq. (6) using the Fourier shifting theorem [23]. Up to here, T. Vallius et al. claimed that this is an efficient computation method compared with the calculation of RCWA to the

entire profile. Once the Fourier coefficient of the step-transition perturbation P_m and P_{-m} is calculated and no further RCWA calculations are necessary.

We furthermore focus on the fact that this Fourier-domain contribution from step transition P_m don't contains explicit dependence on transition point x_k . This point is highly useful when calculating the gradient of diffraction efficiencies with respect to transitions positions

$\nabla_x F = \left[\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_k}, \dots, \frac{\partial F}{\partial x_{2K}} \right]$ to optimize the structures. To find these derivatives, we apply chain rule when differentiating the FOM $F(x)$:

$$\begin{aligned} \frac{\partial F}{\partial x_k} &= \sum_{m=-M}^M \frac{\partial F}{\partial \eta_m} \cdot \frac{\partial \eta_m}{\partial x_k} \\ &= \sum_{m=-M}^M \frac{\partial F}{\partial \eta_m} \cdot \frac{\partial |T_m + D_m|^2}{\partial x_k} \end{aligned} \quad (8)$$

where the first term $\frac{\partial F}{\partial \eta_m}$ is easily calculated by using Eq. (1) and the second term $\frac{\partial |T_m + D_m|^2}{\partial x_k}$ is also expressed by an analytical equation because P_m and P_{-m} don't include the dependence on the position of transition point x_k . Additional details on this are described in the Appendix. It is feasible to calculate the gradient straightforwardly with accuracy as much as the approach based on the rigorous method if most of the features of the structure are bigger than the wavelength of the incident light. The obtained gradient was used in optimization based on the limited-memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) algorithm [24, 25].

3 Simulation results

Using the proposed optimization approach, we can design various multi-spot array generators. In general, diffractive beam splitter creating larger number of spots require more complex structure, *i.e.* gratings with many features. To verify our method is valid in high dimensional optimization problems, we show the optimization results of fan-out grating generating many spots, for instance, 1-to-117 diffractive beam splitter.

To evaluate DOEs, we use two different metrics which are uniformity error (UE) and normalized root-mean-square error (NRMS) σ .

$$UE = \frac{\eta_{\max} - \eta_{\min}}{\eta_{\max} + \eta_{\min}} \quad (9)$$

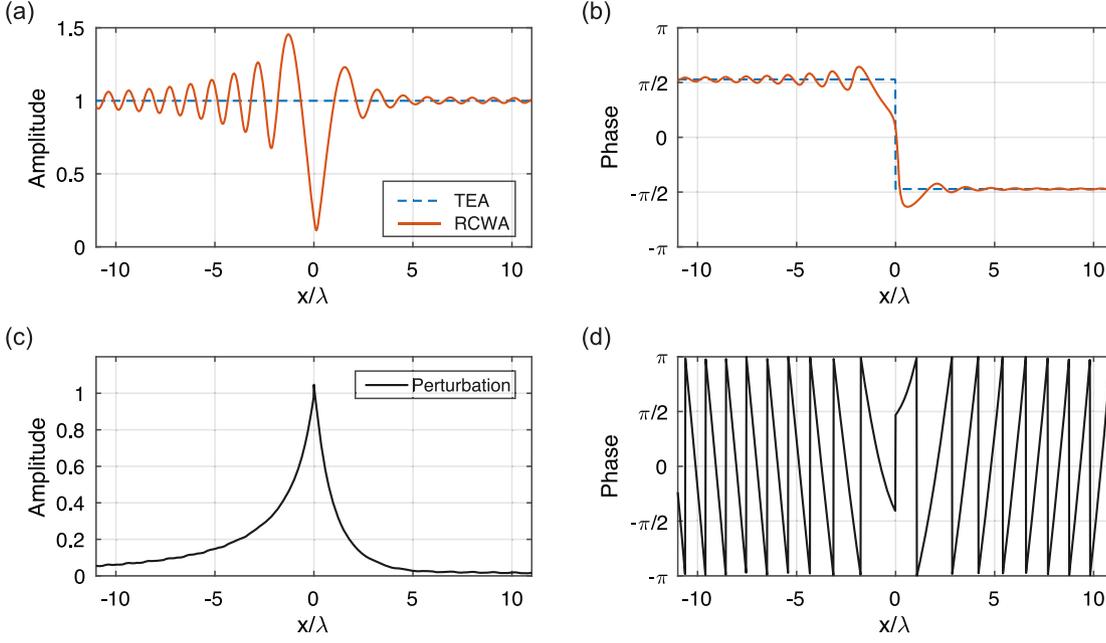


Figure 2: The amplitude (a) and phase (b) of the electric field in TE-polarization and behind the sharp vertical surface profile with a transition point corresponding to a phase delay of π radians calculated by RCWA (red line) and TEA (blue dotted line). The corresponding perturbations in Eq. (2) are shown in (c) and (d).

$$\sigma = \sqrt{\frac{1}{M} \sum \left(\frac{\eta_m - \eta_{obj}}{\eta_{obj}} \right)^2} \quad (10)$$

where η_{\max} and η_{\min} represent the maximal and minimum diffraction intensity and η_m is diffraction efficiency in orders and η_{obj} is target diffraction efficiency and M is the total number of diffraction orders. Lower values of both UE and NRMS indicate less residual variance so that our objective is to minimize UE and NRMS of a DOE design given uniform diffraction efficiency distribution.

To apply the optimization method, we prepared surface profiles of 1D fan-out grating designed by IFTA. The fused silica (SiO_2) was selected as the material. The refractive index of SiO_2 is assumed as $n_2 = 1.46$. Transverse electric (TE)-polarized (*i.e.*, E-field component along the y -axis) monochromatic light with a wavelength of $\lambda = 633$ nm is an incident plane wave from the substrate side with normal incidence angle. The depth of the grating was selected as $d = 692$ nm and the grating period is $200 \mu\text{m}$. Thus, the maximal diffraction angle of 1-to-117 diffractive beam splitter are about 11° at 58th order from 0th order.

To optimize these 1-to-117 diffractive beam splitters, we use our figure of merit as in Eq. (1) with the uniform intensity distribution of target efficiency η_{obj} and find the local optima using the L-BFGS algorithm with the gradient calculated based on STPA. The uniformity of beam array

created by elements designed from IFTA based on TEA, followed by optimization, are plotted in Figure 3 with different metrics. We also plot together with the uniformity of final design after gradient-based optimization based on RCWA, in this case, the gradient calculated by brute-force approaches. In other words, normally around 60 times (*i.e.*, the number of transition points $2K + 1$) RCWA simulation is required in an iteration during the optimization. We compared the optimized results by gradient-based on STPA and RCWA. For an accurate comparison, all diffraction efficiencies of final designs are calculated by RCWA. In most cases, the uniformity of these final elements is significantly improved and the uniformity of final design optimized based on STPA are as good as those of optimized based on RCWA. However, the performance of optimization based on STPA is much better than based on RCWA in terms of computation effort. The simulation and optimization steps were written using MATLAB scripts, and the optimization process took less than 20 s using gradient-based optimization by STPA, while taking over 6 h using gradient-based optimization by RCWA on a machine with 3.60 GHz clock rate and 32 GB RAM. During the optimization, the diffraction pattern for calculating UE and NRMS were evaluated with RCWA solver RETICOLO [26].

To observe the changes of the FOM and transition positions during the optimization, we represent one optimized 1-to-117 diffractive beam splitter in Figure 4.

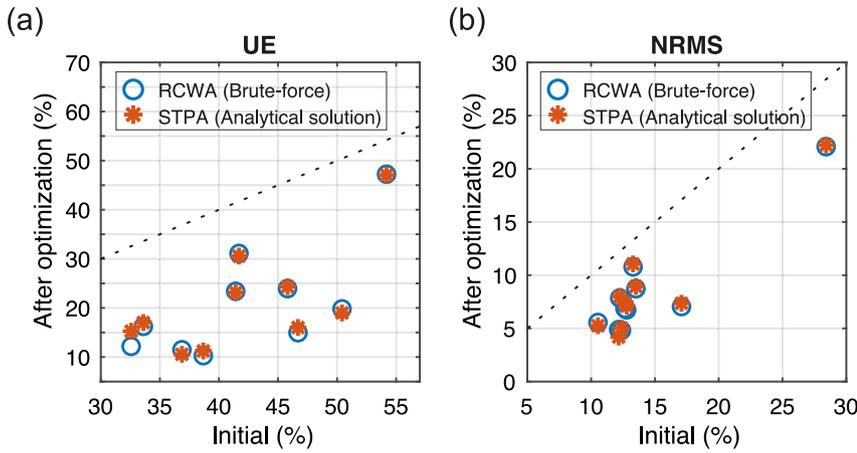


Figure 3: The uniformity of final elements (10 in total) obtained using gradient-based optimization by RCWA and STPA from initial TEA-based IFTA designs with different metrics: (a) UE, (b) UE without zero-th order, and (c) NRMS. The gradient-based on RCWA calculated by Brute-force search. The equal uniformity values of initial and optimized elements are indicated by the black dash line.

Figure 4(a) shows the merit function as a function of the optimization iterations. The figure of merit converged well and the algorithm found the optimum point after 190 iterations. Through the optimization, the change of all transition positions of the structure is plotted in Figure 4(b). The total number of transitions is 66 and the average change of transition points is around 300 nm after optimization. The simulated diffraction efficiency distributions of DOEs after optimization is shown in Figure 4(c). We calculated the total diffraction efficiency, UE, and NRMS of optimized diffractive beam splitters. The total diffraction efficiency of 117 spots of optimized DOE is 77.35% and UE

from 38.68 to 10.79% and NRMS from 12.16 to 04.18%, through gradient-based optimization using STPA. The surface profile of optimized design which has critical dimension (CD) (*i.e.*, minimum feature size) is 700 nm and fill factor is 51.16% is represented in Figure 4(c) inset.

4 Experimental results

The diffractive beam splitters were fabricated by direct laser writing to obtain SiO₂ binary surface relief structures. The elements are optically characterized using a

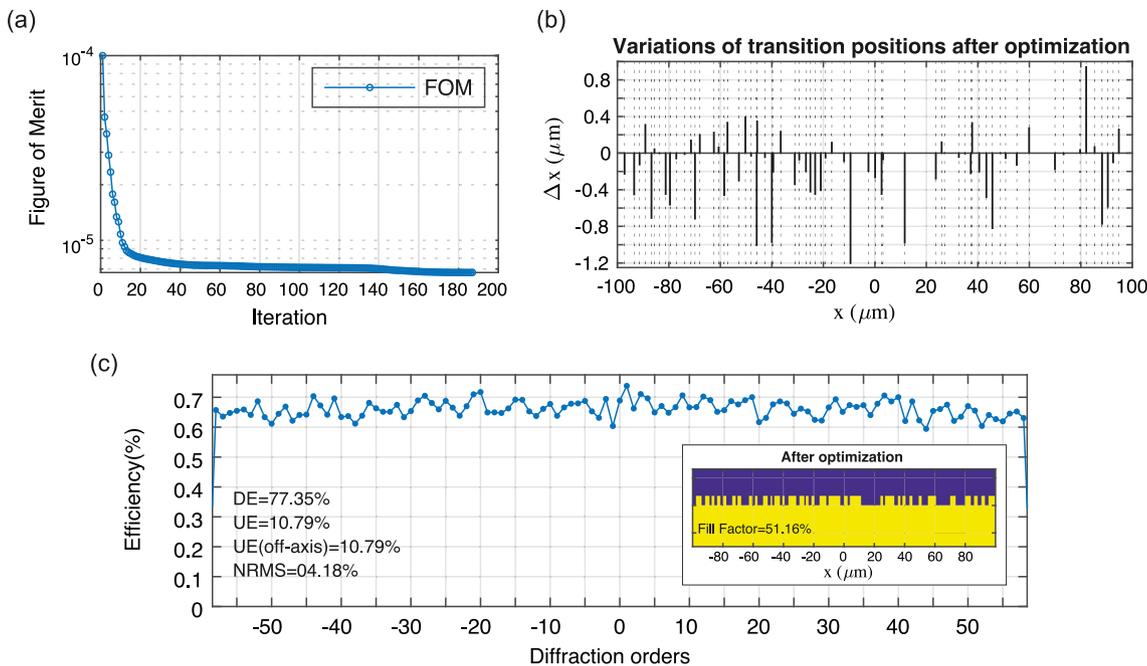


Figure 4: Theoretical analysis of 1-to-117 diffractive beam splitters. (a) the plot of the figure of merit over the course of the optimization process. (b) The variation of transition positions after optimization. The dash lines indicate the transition positions in the single grating period. (c) the calculated efficiency of diffractive beam splitter after optimization. The surface profile of the optimized element in the inset.

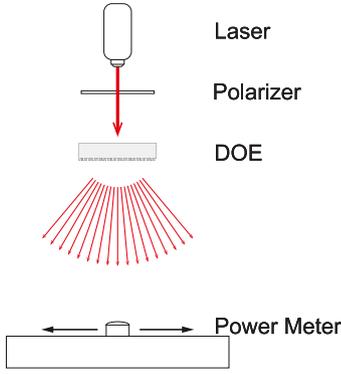


Figure 5: Schematic of equipment used for diffractive array measurements.

TE-polarized 636 nm wavelength beam from a diode laser. We detect the diffracted light beams using a mobile single-pixel detector with a high dynamic range. In Figure 5, a detector with a pinhole aperture is mounted on a translation stage under computer control. By scanning the detector to the center of each of the spots, it is possible to measure the power contained in each of the spots, *i.e.* diffraction orders in the output array.

To focus on both the simulation and experiment to facilitate a quantitative comparison, we applied loss caused by Fresnel reflection from the interface between air

and SiO₂ substrate to simulate the overall efficiency of DOEs. The comparison between theoretical and experimental diffraction efficiencies are presented in Figure 6. We represent the total diffraction efficiency, UE, and NRMS of simulated and measured one in Table 1.

The experimental data show that the DOE operates with high-performance. The UE and NRMS of beam splitters are 21.42 and 8.07%, respectively. For an accurate comparison between theoretical and measured results, we analyze the correlation of these data using mean absolute percentage deviation (MAPD) as a ratio defined by the formula:

$$\text{MAPD} = \frac{1}{M} \sum \left| \frac{\eta_m^S - \eta_m^E}{\eta_m^S} \right| \quad (11)$$

Table 1: Comparison with the simulated and experimental properties of the 1-to-117 beam splitters. The simulated efficiency take into account the loss from Fresnel reflection in the air-SiO₂ substrate interface.

	Simulated	Measured
Total efficiency (%)	74.56	74.65
UE (%)	11.74	21.42
NRMS (%)	04.45	08.07

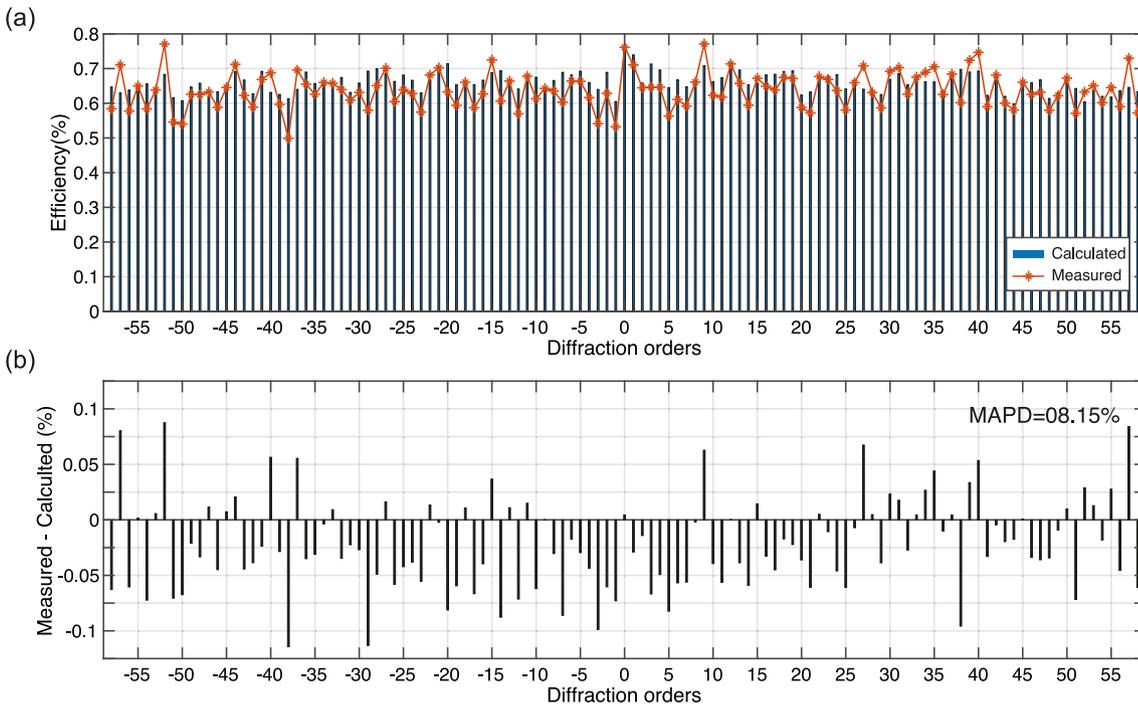


Figure 6: Experimental characterization of 1-to-117 diffractive beam splitter. (a) experimental data (orange star) from profile optimized base on STPA and the simulated data (blue bar). (b) difference between experimental and simulated data in orders.

where η_m^S, η_m^E are simulated and experimental efficiency in (m)th diffraction orders and M is the total number of diffraction orders. The MAPD of 1-to-117 beam splitters are calculated to the 8.15%, which shows excellent reproducibility of the simulated results in a quantitative manner. The only noticeable deviation in the measurement is a small mismatch of diffraction efficiency in a few orders due to minor fabrication errors. In general the diffraction efficiency in orders often strongly depends on the errors in fabrication processes, *e.g.*, etching depth, feature width, slope steepness, and feature rounding. Nevertheless, the fabricated samples based on optimized design overall display experimental performances which are better than the theoretical performances of initial designs before optimization.

5 Conclusion

In summary, we utilized the STPA in optimizing the optical elements, which is able to create wide angle diffractive optical elements at a very low computational cost. We explored properties of the optimization method, such as efficient computation for the gradient of the target function with respect to transition positions with Fourier-domain local field perturbation. As a case study, we applied gradient-based optimization with STPA to 1–117 beam splitter with a non-paraxial diffraction angle, *i.e.*, maximal diffraction angle is 11° from the center, respectively. The optimized beam splitter show a considerable improvement of uniformity while maintaining the initial diffraction efficiency. The experimental results obtained by the illumination of the fabricated optical elements using a laser of 635 nm wavelength with a normal incidence have been compared with the numerical results. Numerical simulation and experimental results were found to be in good agreement and our optimization method can be considered proven to be an effective design tool for wide angle diffractive beam splitters.

Author contribution: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: Horizon 2020 Framework Programme (675745).

Conflict of interest statement: The authors declare no conflicts of interest regarding this article.

Appendix

In order to calculate the gradient of figure of merit in Eq. (8), we calculated the derivatives of the diffraction efficiencies $\frac{\partial \eta_m}{\partial x_k}$.

$$\frac{\partial \eta_m}{\partial x_k} = \frac{\partial |T_m|^2}{\partial x_k} + \frac{\partial |D_m|^2}{\partial x_k} + \frac{\partial T_m^* D_m}{\partial x_k} + \frac{\partial T_m D_m^*}{\partial x_k} \quad (12)$$

where the diffraction efficiency η_m is a function with respect to transition point x_k in binary grating. Thus we can partially differentiate each term of η_m with respect to x_k . when $m \neq 0$, we can express the derivatives as

$$\frac{\partial |T_m|^2}{\partial x_k} = 2\Phi_m^2 (C_{1m} C'_{1m} + S_{1m} S'_{1m}) \quad (13a)$$

$$\frac{\partial |D_m|^2}{\partial x_k} = \begin{cases} |P_m|^2 (C_{2m} C'_m + S_{2m} S'_m) \\ + 2\Re(P_m P_{-m}^*) (C_{3m} C'_m + S_{3m} S'_m) \\ - 2\Im(P_m P_{-m}^*) (S_{3m} C'_m - C_{3m} S'_m), \\ \quad \text{for } k = 1, 3, \dots, 2K - 1 \\ |P_{-m}|^2 (C_{3m} C'_m + S_{3m} S'_m) \\ + 2\Re(P_m P_{-m}^*) (C_{2m} C'_m + S_{2m} S'_m) \\ - 2\Im(P_m P_{-m}^*) (C_{2m} S'_m - S_{2m} C'_m), \\ \quad \text{for } k = 2, 4, \dots, 2K \end{cases} \quad (13b)$$

$$\frac{\partial T_m^* D_m}{\partial x_k} + \frac{\partial T_m D_m^*}{\partial x_k} = \begin{cases} 2\Phi_m \left[\Re(P_m) (C_{1m} C'_{1m} + C_{2m} C'_{1m} + S_{1m} S'_{1m} + S_{2m} S'_{1m}) \right. \\ \quad + \Im(P_m) (C_{1m} S'_{1m} - S_{1m} C'_{1m} - C_{2m} S'_{1m} + S_{2m} C'_{1m}) \\ \quad + \Re(P_{-m}) (C_{3m} C'_{1m} + S_{3m} S'_{1m}) \\ \quad \left. + \Im(P_{-m}) (S_{3m} C'_{1m} - C_{3m} S'_{1m}) \right], \\ \quad \text{for } k = 1, 3, \dots, 2K - 1 \\ 2\Phi_m \left[\Re(P_{-m}) (C_{3m} C'_{1m} + C_{1m} C'_m + S_{3m} S'_{1m} + S_{1m} S'_m) \right. \\ \quad + \Im(P_{-m}) (S_{3m} C'_{1m} + C_{1m} S'_m - C_{3m} S'_{1m} - S_{1m} C'_m) \\ \quad + \Re(P_m) (C_{2m} C'_{1m} + S_{2m} S'_{1m}) \\ \quad \left. + \Im(P_m) (S_{2m} C'_{1m} - C_{2m} S'_{1m}) \right], \\ \quad \text{for } k = 2, 4, \dots, 2K \end{cases} \quad (13c)$$

where

$$\Phi_m = \sin(\Delta\phi/2)/\pi m \quad (14a)$$

$$C_{1m} = \sum_{k=1}^{2K} (-1)^k \cos(2\pi m x_k / \Lambda) \quad (14b)$$

$$S_{1m} = \sum_{k=1}^{2K} (-1)^k \sin(2\pi m x_k / \Lambda) \quad (14c)$$

$$C_{2m} = \sum_{k=1}^K \cos(2\pi m x_{2k-1} / \Lambda) \quad (14d)$$

$$S_{2m} = \sum_{k=1}^K \sin(2\pi m x_{2k-1} / \Lambda) \quad (14e)$$

$$C_{3m} = \sum_{k=1}^K \cos(2\pi m x_{2k} / \Lambda) \quad (14f)$$

$$S_{3m} = \sum_{k=1}^K \sin(2\pi m x_{2k} / \Lambda) \quad (14g)$$

$$C'_{1m} = \frac{2\pi m}{\Lambda} (-1)^k \sin(2\pi m x_k / \Lambda) \quad (14h)$$

$$S'_{1m} = \frac{2\pi m}{\Lambda} (-1)^k \cos(2\pi m x_k / \Lambda) \quad (14i)$$

$$C'_m = \frac{2\pi m}{\Lambda} \sin(2\pi m x_k / \Lambda) \quad (14j)$$

$$S'_m = \frac{2\pi m}{\Lambda} \cos(2\pi m x_k / \Lambda). \quad (14k)$$

The $\Delta\phi$ is the difference between phase ϕ_1 and ϕ_2 which are the phase of an electric field in the air and dielectric material, respectively and $2K$ is the number of transition point in structure. The Fourier coefficients P_m and P_{-m} of field perturbation are given by Eq. (7), which are constant values with respect to transition point x_k . Thus the values $\Re(P_m)$, $\Re(P_{-m})$, $\Im(P_m)$, $\Im(P_{-m})$, $\Re(P_m P_{-m}^*)$ and $\Im(P_m P_{-m}^*)$ are constant with respect to transition point x_k .

If $m = 0$, the derivatives of the diffraction efficiency in zero order is expressed as

$$\begin{aligned} \frac{\partial \eta_0}{\partial x_k} = & -4Q'(1-2Q) \sin^2(\Delta\phi/2) \\ & -8K \cdot \Re(P_m) \sin(\Delta\phi/2) \sin(\phi_s/2) Q' \\ & + 8K \cdot \Im(P_m) \sin(\Delta\phi/2) \cos(\phi_s/2) Q' \end{aligned} \quad (15)$$

where $Q = \sum_{k=1}^{2K} (-1)^k x_k$, $Q' = \frac{(-1)^k}{\Lambda}$, and $\phi_s = \phi_1 + \phi_2$. Therefore, we can express the gradient of diffraction efficiency with respect to transition points based on STPA as an analytical solution.

References

- [1] J. E. Jureller, H. Y. Kim, and N. F. Scherer, "Stochastic scanning multiphoton multifocal microscopy," *Opt. Express*, vol. 14, p. 3406, 2006.
- [2] Z. Chen, B. Mc Larney, R. Johannes, et al, "High-speed large-field multifocal illumination fluorescence microscopy," *Laser Photonics Rev.*, vol. 14, 2020, Art no. 1900070.
- [3] M. Bauer, D. Griessbach, A. Hermerschmidt, S. Krüger, M. Scheele, and A. Schischmanow, "Geometrical camera calibration with diffractive optical elements," *Opt. Express*, vol. 16, pp. 20241–20248, 2008.
- [4] F. Wang, Z. Zhang, R. Wang, et al, "Distortion measurement of optical system using phase diffractive beam splitter," *Opt. Express*, vol. 27, pp. 29803–29816, 2019.
- [5] D. F. Brosseau, F. Lacroix, M. H. Ayliffe, et al, "Design, implementation, and characterization of a kinematically aligned, cascaded spot- array generator for a modulator-based free-space optical interconnect," *Appl. Opt.*, vol. 39, p. 733, 2000.
- [6] R. Vandenhousten, A. Hermerschmidt, and R. Fiebelkorn, "Design and quality metrics of point patterns for coded structured light illumination with diffractive optical elements in optical 3D sensors," in *Digital Optical Technologies International Society for Optics and Photonics 2017*, B. C. Kress and P. Schelkens, vol. 10335, 2017, pp. 264–276, SPIE.
- [7] O. Barlev and M. A. Golub, "Multifunctional binary diffractive optical elements for structured light projectors," *Opt. Express*, vol. 26, p. 21092, 2018.
- [8] T. Simon, A. Arfaoui, and P. Desaulniers, "Cross-diffractive optical elements for wide angle geometric camera calibration," *Opt. Lett.*, vol. 36, p. 4770, 2011.
- [9] P. Twardowski, B. Serio, V. Raulot, and M. Guilhem, "Three-dimensional shape measurement based on light patterns projection using diffractive optical elements," in *Micro-Optics 2010*, H. Thienpont, P. Van Daele, J. Mohr, and H. Zappe, Eds., vol. 7716, International Society for Optics and Photonics, 2010, pp. 704–711, SPIE.
- [10] W. Frank and O. Bryngdahl, "Iterative fourier-transform algorithm applied to computer holography," *J. Opt. Soc. Am. A*, vol. 5, p. 1058, 1988.
- [11] M. Skeren, I. Richter, and P. Fiala, "Iterative fourier transform algorithm: comparison of various approaches," *J. Mod. Opt.*, vol. 49, pp. 1851–1870, 2002.
- [12] S. Bühling and W. Frank, "Improved transmission design algorithms by utilizing variable-strength projections," *J. Mod. Opt.*, vol. 49, pp. 1871–1892, 2002.
- [13] J. W. Goodman, "Introduction to fourier optics," in *Introduction to Fourier Optics*, vol. 1, 3rd ed., J. W. Goodman, Ed., Englewood, CO, Roberts & Co. Publishers, 2005 Chapter 5. p.
- [14] D. A. Pomet, M. G. Moharam, and E. B. Grann, "Limits of scalar diffraction theory for diffractive phase elements," *J. Opt. Soc. Am. A*, vol. 11, p. 1827, 1994.
- [15] T. Vallius, P. Vahimaa, and M. Honkanen, "Electromagnetic approach to the thin element approximation," *J. Mod. Opt.*, vol. 51, pp. 2079–2092, 2004.
- [16] P. Lalanne and G. M. Morris, "Highly improved convergence of the coupled-wave method for TM polarization," *J. Opt. Soc. Am. A*, vol. 13, p. 779, 1996.
- [17] M. G. Moharam, E. B. Grann, D. A. Pomet, and T. K. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled-wave analysis of binary gratings," *J. Opt. Soc. Am. A*, vol. 12, pp. 1068–1076, 1995.
- [18] V. Liu and S. Fan, "S 4: A free electromagnetic solver for layered periodic structures," *Comput. Phys. Commun.*, vol. 183, pp. 2233–2244, 2012.
- [19] S. Ruder, *An overview of gradient descent optimization algorithms*, CoRR abs/1609.04747, 2016.

- [20] T. Vallius, V. Kettunen, M. Kuittinen, and J. Turunen, "Step-discontinuity approach for non-paraxial diffractive optics," *J. Mod. Opt.*, vol. 48, pp. 1195–1210, 2001.
- [21] F. J. Wen, and P. S. Chung, "2D optical beam splitter using diffractive optical elements (DOE)," in *Passive Components and Fiber-based Devices III*, Sang Bae Lee, Yan Sun, Kun Qiu, Simon C Fleming and IanWhite Ed., vol. 6351, International Society for Optics and Photonics, 2006, pp. 200–210, SPIE.
- [22] S. Banerji and B. Sensale-Rodriguez, "A computational design framework for efficient, fabrication error-tolerant, planar THz diffractive optical elements," *Sci. Rep.*, vol. 9, p. 5801, 2019.
- [23] R. L. Easton, *Fourier Methods in Imaging*, John Wiley & Sons, Ltd, Apr. 2010. <https://doi.org/10.1002/9780470660102>.
- [24] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu, "A limited memory algorithm for bound constrained optimization," *SIAM J. Sci. Comput.*, vol. 16, pp. 1190–1208, 1995.
- [25] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal, "Algorithm 778: L-BFGS-B: fortran subroutines for large-scale bound-constrained optimization," *ACM Trans. Math. Softw.*, vol. 23, pp. 550–560, 1997.
- [26] J.-P. Hugonin and P. Lalanne, *Reticolo Software for Grating Analysis*, Orsay, France, Institut d'Optique, 2005.