Optics Letters

Effects of interlayer reflection and interpixel interaction in diffractive optical neural networks

Minhan Lou,[1] Yingjie Li,[1] Cunxi Yu,[1,2] Berardi Sensale-Rodriguez,[1,3] and Weilu Gao[1,*]

[1]Department of Electrical and Computer Engineering, The University of Utah, 50 S Central Campus Drive, Utah 84112, USA
[2]e-mail: cunxi.yu@utah.edu
[3]e-mail: berardi.sensale@utah.edu
*Corresponding author: weilu.gao@utah.edu

Received 10 October 2022; revised 17 November 2022; accepted 3 December 2022; posted 5 December 2022; published 2 January 2023

Multilayer diffractive optical neural networks (DONNs) can perform machine learning (ML) tasks at the speed of light with low energy consumption. Decreasing the number of diffractive layers can reduce inevitable material and diffraction losses to improve system performance, and incorporating compact devices can reduce the system footprint. However, current analytical DONN models cannot accurately describe such physical systems. Here we show the ever-ignored effects of interlayer reflection and interpixel interaction on the deployment performance of DONNs through full-wave electromagnetic simulations and terahertz (THz) experiments. We demonstrate that the drop of handwritten digit classification accuracy due to reflection is negligible with conventional low-index THz polymer materials, while it can be substantial with high-index materials. We further show that one- and few-layer DONN systems can achieve high classification accuracy, but there is a trade-off between accuracy and model-system matching rate because of the fast-varying spatial distribution of optical responses in diffractive masks. Deep DONNs can break down such a trade-off because of reduced mask spatial complexity. Our results suggest that new accurate and trainable DONN models are needed to advance the development and deployment of compact DONN systems for sophisticated ML tasks.

https://doi.org/10.1364/OL.477605

Machine learning (ML) algorithms have been widely employed in diverse applications, such as computer vision and scientific computing [1,2]. A recent discovery by Lin *et al.* offers new opportunities of manipulating the wavefront of terahertz (THz) beams in diffractive optical components to perform ML tasks in a high-throughput and energy-efficient manner [3]. The cascaded multiple diffractive masks, which are coined as diffractive optical neural networks (DONNs), can accurately classify handwritten digits in the MNIST dataset. Since then, extensive numerical and experimental studies of DONNs [4] have extended their ML capability, such as the classification of more complex datasets [5] and multi-task classification [6,7], and have enabled new hardware implementation with compact metasurfaces in readily accessible visible and near-infrared wavelengths [8–12] and reconfigurable diffractive masks using spatial light modulators [13–15].

The first demonstrated diffractive masks were prototyped through 3D printing in a low-index polymer [3] in the THz range. Furthermore, in the visible ranges, advanced nanophotonic and deep subwavelength structures with high-index materials have been employed for compact integration [7,16]. However, the assumption of treating each diffraction pixel as an infinitely large slab for calculating the optical response in current analytical DONN models breaks down when the pixel size becomes comparable to the wavelength and efficient diffraction creates a fast-varying incident field in the subwavelength scale [17]. The resulting discrepancy between models and experiments has been reported in the design of subwavelength structures [18,19]. Furthermore, in previous experimental demonstrations of free-space DONNs for 10-label MNIST classification, there are an observable accuracy drop and field distribution mismatch between experimental results and model calculations, which are not explained or generally attributed to experimental imperfection [3,9,20]. The interpixel interaction in diffractive masks have been proposed as one contributing factor to the model-experiment mismatch in an on-chip 2-label DONN classifier [10], while there is no consensus in the other free-space 2-label DONN classifier [7]. Moreover, the effects of multiple reflection between diffractive masks have not been considered in DONN models before.

Here, we demonstrate an analysis of the effects of interlayer reflection and interpixel interaction in THz DONNs on the observed discrepancy between experiments and model calculations. We show that the reflection effect is negligible with low-index materials and is substantial with high-index materials, which are beneficial for DONNs. For example, silicon with THz index ~ 3.4 can reduce the diffractive mask thickness and enable small-volume resonant meta-atoms for compact DONNs. In addition, emerging chalcogenide phase-change materials with THz indices >10 [21] can enable reconfigurable DONNs. Moreover, we employ full-wave finite-domain time-difference (FDTD) numerical simulations to obtain the classification accuracy, which agrees better with experiments than the analytical

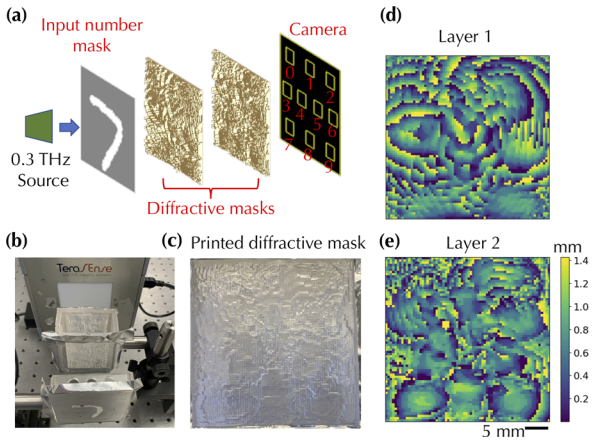

Fig. 1. Overview of a two-layer THz DONN system, model, and components. (a) Diagram of the two-layer THz DONN system. (b) A number mask (digit 7), two phase masks, and a THz camera used in experimental setup. (c) A close look at one printed mask. The height map of (d) the first and (e) the second layer of trained masks.

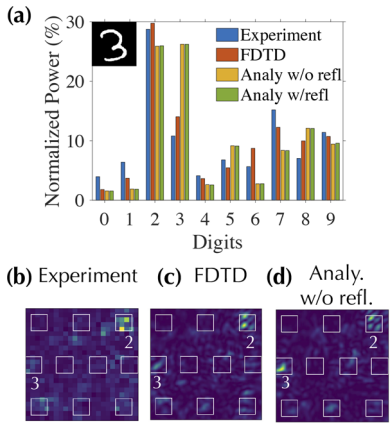

Fig. 2. Experimental and simulation results of the two-layer DONNs system. (a) Power ratio in each detector region obtained from experiments, FDTD simulation, and the analytical models with and without reflection effect, as well as (b) experimentally measured, (c) FDTD calculated, (d) analytical model (without reflection) calculated camera output images for an input digit 3.

model. We discover that the fast spatial variation of complex-valued transmission coefficients in diffractive masks of the two-layer DONN system is needed to obtain high classification accuracy, which leads to the substantial mismatch between FDTD simulations and analytical models. Increasing the DONN depth reduces the mask spatial variation for high classification accuracy and is advantageous in precisely deploying trained analytical DONN models to experimental systems. Our results suggest that the development of accurate trainable DONN simulators will be crucial for the design and deployment of compact DONN systems to tackle complex ML tasks.

Figure 1(a) shows the diagram of a two-layer THz DONN setup. A 0.3-THz (1-mm wavelength) beam generated from a Virginia Diodes TX267 transmitter with a power ~ 25 mW is incident onto a number mask, which is followed by two diffractive masks. Diffractive masks are fabricated using an ELEGOO Saturn Resin 3D printer and their edges are covered by an aluminum film. The mask size is 3.6 cm × 3.6 cm with a step resolution of 0.5 mm, which means 72 × 72 pixels. The refractive index of the 3D printing resin polymer is estimated to be $1.7 + 0.03j$. The number mask is fabricated by first 3D printing the digit template and then transferring the digit patten onto an aluminum film. Specifically, the dark region of the digit is covered by the aluminum film and the bright region is open by cutting the aluminum film with a razor blade. For the digit with disconnected areas, such as 0, the bright area in the 3D-printed template is a thin film. A Tera-1024 THz camera from Terasense Group is used to capture the output images of the DONN system. Ten pre-defined regions in the THz camera represent ten digits in the MNIST dataset. The distance between the number mask, diffractive masks, and the camera is the same as 3 cm, which is selected for good DONN performance and manageable simulation memory. Figure 1(b) displays the experimental setup, where two diffractive masks are printed as a whole to minimize misalignment. Figure 1(c) shows a close-look of a 3D-printed diffractive mask. Figures 1(d) and 1(e) show the spatial height mapping of two trained diffractive masks to perform 10-label MNIST classification.

In the analytical DONN model for training, each diffractive pixel in diffractive masks is treated as a cuboid with a square base. The lateral length of the square base is fixed and defined as the pixel size, and the cuboid height is a trainable parameter. The transmission electric field is a product of the incident field and the field transmission coefficient at each pixel location. The input and output planes for calculating transmission coefficients are defined as two wavefront planes with a 2π phase difference. For the approximation, the field transmission coefficient at a specific pixel location is calculated as the transmission coefficient of an infinitely extended slab with thickness equal to the cuboid height at that location. Thus, based on the transfer matrix method (see Supplement 1, Note 1), the electric field transmission coefficient $t_{i,n}$ on ith pixel in the nth diffractive mask is a function of the height $h_{i,n}$ and refractive index n_m of that pixel, which is calculated as

$$t_{i,n} = \frac{4n_m \exp(j(n_m - 1)k_0 h_{i,n})}{(1 + n_m)^2 - (1 - n_m)^2 \exp(j2n_m k_0 h_{i,n})}.$$

The Rayleigh–Sommerfeld scalar free-space diffraction function and fast Fourier transform are used to calculate the electric field propagation between diffractive masks [3,6,14]. To meet the Nyquist sampling criterion of the Rayleigh–Sommerfeld diffraction [22] and balance memory resource requirement, each physical diffractive pixel is divided into a group of 4 × 4 subpixels that have the same optical response but different incident field. Thus, the calculation pixel size is 0.125 mm.

Since it is challenging to fabricate grays-level number masks in experiment, input images from the MNIST dataset are binarized with a threshold at 0.5 and scaled up to match the size of diffractive masks. To account for the input Gaussian THz beam, the input field to the DONN system is a product of a binarized number mask and a normalized 2D Gaussian distribution. On the camera plane, the softmax activation function is employed to calculate the loss function based on the camera reading.

Figure 2 shows a substantial discrepancy between the light power in pre-defined detector regions obtained from analytical model calculations and that in the experiment. For an example input digit 3 in Fig. 2(a), the power ratio at target region 3 drops significantly from 26.2% in the analytical model without incorporating the reflection effect to a value of 10.8% in the experiment. The power ratio at non-target region 2 increases

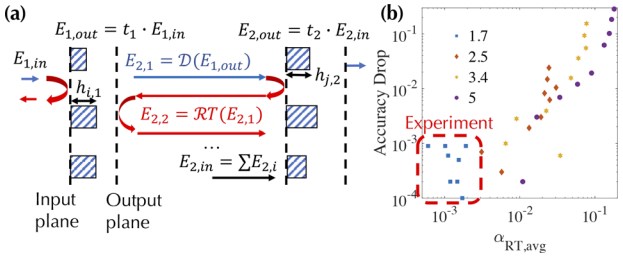

Fig. 3. Interlayer reflection effect. (a) Schematic of incorporating interlayer reflection effect. (b) Accuracy drop as a function of average round trip power ratio $\alpha_{RT,\,avg}$ for multiple trained diffractive masks with different refractive indices.

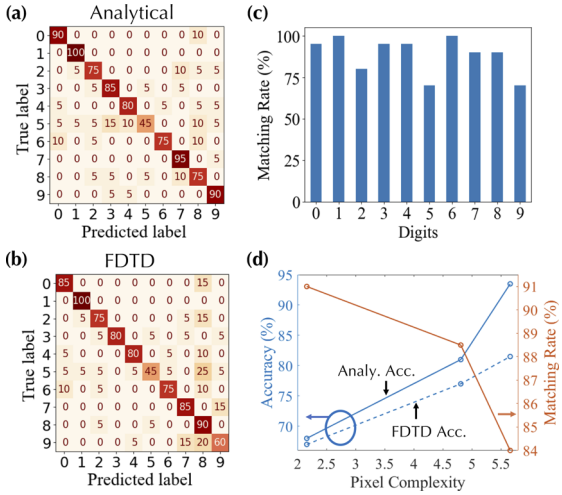

Fig. 4. Interpixel interaction effect. Confusion matrices of classifying 200 MNIST images calculated using (a) the analytical model and (b) FDTD simulation. (c) Matching rate between the analytical model and FDTD simulation for images of each digit. (d) Classification accuracy obtained from analytical and FDTD approaches and their matching rate as a function of the complexity of diffractive masks.

from 25.8% to 28.7%. As the result, the experimental classification becomes wrong as digit 2; see Fig. 2(b). The experimental results agree well with FDTD simulations (see Supplement 1, Note 2) in Fig. 2(c), and both show brighter spots in region 2 than region 3. In contrast, these results deviate from the analytical model without considering the reflection effect, which displays a brighter spot at the region 3, as shown in Fig. 2(d). Furthermore, for a correctly classified digit, the power ratio at the target region in the experiment and FDTD simulations is substantially smaller than that in the analytical mode (see Supplement 1, Note 3).

We first evaluate the effect of interlayer reflection on the observed discrepancy between the analytical model and experiments. Figure 3(a) illustrates the reflection between layers and the diffraction of the reflected field. Here, $\mathcal{D}$ represents the diffraction operation on the field, and $\mathcal{RT}$ represents the round trip operation that consists of the reflection on both diffractive masks and two diffraction operations. The field reflection coefficients can also be calculated using the transfer matrix method. The field associated with multiple reflection–diffraction round trips can be iteratively included in calculating the steady-state incident field onto the second diffractive mask (see Supplement 1, Note 4). The DONN system is first trained and validated using the analytical model without reflection effect. The classification accuracy of the obtained trained system is then evaluated on randomly shuffled 10,000 images from the MNIST dataset using the analytical model incorporating interlayer reflection. The drop of classification accuracy is used to evaluate the reflection effect. A few materials with the real parts of refractive indices 1.7, 2.5, 3.4, and 5 are selected for analysis. The imaginary part is fixed as 0.03. Since the power of the reflected and following diffracted light is not only dependent on the material index but also on diffractive masks, we train multiple diffractive masks for a given material index. For a given input image, we define a round trip power ratio α_{RT} based on described fields in Fig. 3(a) as

$$\alpha_{RT} = \frac{\iint |\mathcal{RT}(E_{2,1}(x,y))|^2\,dx\,dy}{\iint |E_{2,1}(x,y)|^2\,dx\,dy}.$$

Since α_{RT} is also dependent on the incident field distribution, we then take the average of α_{RT} over 10,000 input images, which is denoted as $\alpha_{RT,\,avg}$. Each digit has 1000 input images. Figure 3(b) displays the accuracy drop as a function of $\alpha_{RT,\,avg}$ for different trained diffractive masks with different material indices. A clear correlation between increasing $\alpha_{RT,\,avg}$ and increasing accuracy drop is observed. In the THz DONN system, the 3D printing polymer has a refractive index 1.7 and the overall accuracy drop is negligible [<0.1%, dashed rectangle in Fig. 3(b)]. Furthermore, there is no noticeable difference in the light power distribution on output images calculated from the analytical models with and without the reflection effect; see Fig. 2(a). In contrast, the diffractive masks fabricated on the material of refractive index 5 can lead to ~ 30% accuracy drop because of a large $\alpha_{RT} \sim 18\%$. These results suggest that for high-index materials, the reflection effect needs to be incorporated in the trainable DONN system model for accurate hardware deployment.

Instead of the interlayer reflection effect, the interpixel interaction effect can lead to the observed discrepancy between experiments and analytical model. We evaluate such an effect by comparing the classification accuracy obtained from the analytical model with that from FDTD, which can capture all near-field and far-field interactions and precisely describe experiments [18]. We randomly choose 200 input images with 20 images for each digit to simulate the output images of a trained DONN system using cascaded FDTD calculations. For example, a given set of trained diffractive masks using the analytical model has a classification accuracy of 81.0% for the 200 input images; see the confusion matrix in Fig. 4(a). The FDTD calculation of the same set of diffractive masks and input images yields a classification accuracy of 77.5%; see the confusion matrix in Fig. 4(b). In addition to the accuracy drop, the classification results of each digit label from two calculation approaches show various degrees of discrepancy. For example, both correct and wrong classification of digits 1 and 6 demonstrate a perfect match while the matching rate for digit 9 is low. Despite the lack of accuracy drop for digits 2 and 5, some wrong and correct classifications of input images become flipped in FDTD simulations and thus the match rate is also low. In contrast, digit 8 even shows an increased accuracy. Figure 4(c) displays the matching rate for all digits.

We further train multiple diffractive masks with different classification accuracies. The best accuracy we can obtain with two 72×72 diffractive masks is 93.5%, which is already better than the first demonstration with five 200×200 diffractive masks

and 91.8% accuracy [3]. We discover that trained diffractive masks have to be quite complex for high-accuracy classification; see the solid blue line in Fig. 4(d) and Supplement 1, Note 5. Quantitatively, we define the average pixel complexity C for a DONN system with L layers of $N \times M$ masks with pixel length of p as

$$C = \sum_{l=1}^{L}\sum_{n=1}^{N}\sum_{m=1}^{M} \frac{\lambda}{LNMp} C_{n,m,l}, \text{where } C_{n,m,l} = |t_{n,m,l} - t_{n-1,m,l}|^2 + |t_{n,m,l} - t_{n+1,m,l}|^2 + |t_{n,m,l} - t_{n,m-1,l}|^2 + |t_{n,m,l} - t_{n,m+1,l}|^2,$$
$$\forall_{n,m,l}\ t_{0,m,l} = t_{N+1,m,l} = t_{n,0,l} = t_{n,M+1,l} = 0.$$

Here, t, λ are the field transmission coefficient and the vacuum wavelength, respectively. The maximum value of C is 32 for $\lambda = 1$ mm and $p = 0.5$ mm. However, the diffractive masks with large C lead to a substantial accuracy drop when the trained model is calculated using FDTD simulations, because the optical response of a pixel is considerably affected by neighboring pixels and deviates from infinitely extended slab assumption in the analytical model. Thus, Fig. 4(d) shows a trade-off between the classification accuracy and the matching rate.

Finally, we compare the influence of interlayer pixels on the performance of DONN systems with different depth and fixed component distance of 3 cm. Specifically, we train one-layer, three-layer, and five-layer DONNs all with classification accuracy ~95% (see Supplement 1, Note 6). Strikingly, as shown in Fig. 5(a), a single diffractive mask can achieve ~95% accuracy, although at the expense of large complexity C and low matching rate of 79%. However, the increasing depth of DONNs helps to reduce the complexity of diffractive masks, so that the obtained accuracy from FDTD calculations is closer to that from the analytical model with a larger matching rate; see Fig. 5(b). Specifically, for the five-layer DONNs, the classification accuracy from the analytical model is 96.5% and that from FDTD calculation is 94.0% thanks to the low diffractive mask complexity of 2.26; see confusion matrices in Figs. 5(c) and 5(d). The overall matching rate is as high as 96%. The required complexity to achieve the diffraction capability to direct beams into different detector regions based on input images can be distributed into multiple diffractive masks in deep DONNs, so that each mask complexity is reduced. In addition, gray-level input images with smoother profiles than binary images could also help to reduce mask complexity (see Supplement 1, Note 7).

In summary, we demonstrate that the interlayer reflection effect is substantial when diffractive masks are made of high-index materials. Furthermore, we demonstrate that one- and few-layer DONNs can achieve high-accuracy classification under the analytical model, while the required large spatial complexity in diffractive masks leads to the degraded performance in practical system deployment. Deep DONNs are advantageous to achieve both high accuracy and accurate deployment. Since DONNs are linear systems, our results are also applicable to DONNs in other wavelengths. Thus, fast-computed, trainable, and physics-accurate DONN simulators will be needed for designing and deploying compact DONNs.

Funding. National Science Foundation (1936729, 2019336).

Disclosures. The authors declare no conflicts of interest.

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

Supplemental document. See Supplement 1 for supporting content.

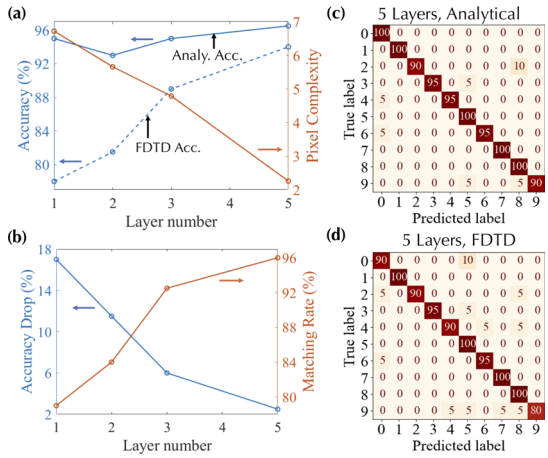

Fig. 5. Advantage of deep DONNs. (a) Classification accuracy obtained from the analytical model and FDTD simulations, diffractive mask complexity, (b) accuracy drop, and matching rate as a function of depth. Confusion matrices for a five-layer DONNs system calculated using (c) the analytical model and (d) FDTD simulations.

REFERENCES

1. Y. LeCun, Y. Bengio, and G. Hinton, Nature **521**, 436 (2015).
2. K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, Nature **559**, 547 (2018).
3. X. Lin, Y. Rivenson, N. T. Yardimci, M. Veli, Y. Luo, M. Jarrahi, and A. Ozcan, Science **361**, 1004 (2018).
4. D. Mengu, M. S. Sakib Rahman, Y. Luo, J. Li, O. Kulce, and A. Ozcan, Adv. Opt. Photonics **14**, 209 (2022).
5. M. S. S. Rahman, J. Li, D. Mengu, Y. Rivenson, and A. Ozcan, Light: Sci. Appl. **10**, 14 (2021).
6. Y. Li, R. Chen, B. Sensale-Rodriguez, W. Gao, and C. Yu, Sci. Rep. **11**, 11013 (2021).
7. X. Luo, Y. Hu, X. Ou, X. Li, J. Lai, N. Liu, X. Cheng, A. Pan, and H. Duan, Light: Sci. Appl. **11**, 158 (2022).
8. E. Goi, X. Chen, Q. Zhang, B. P. Cumming, S. Schoenhardt, H. Luan, and M. Gu, Light: Sci. Appl. **10**, 40 (2021).
9. H. Chen, J. Feng, M. Jiang, Y. Wang, J. Lin, J. Tan, and P. Jin, Engineering **7**, 1483 (2021).
10. T. Fu, Y. Zang, H. Huang, Z. Du, C. Hu, M. Chen, S. Yang, and H. Chen, Opt. Express **29**, 31924 (2021).
11. Z. Wang, L. Chang, F. Wang, T. Li, and T. Gu, Nat. Commun. **13**, 2131 (2022).
12. H. H. Zhu, J. Zou, H. Zhang, Y. Z. Shi, S. B. Luo, N. Wang, H. Cai, L. X. Wan, B. Wang, X. D. Jiang, J. Thompson, X. S. Luo, X. H. Zhou, L. M. Xiao, W. Huang, L. Patrick, M. Gu, L. C. Kwek, and A. Q. Liu, Nat. Commun. **13**, 1044 (2022).
13. T. Zhou, X. Lin, J. Wu, Y. Chen, H. Xie, Y. Li, J. Fan, H. Wu, L. Fang, and Q. Dai, Nat. Photonics **15**, 367 (2021).
14. R. Chen, Y. Li, M. Lou, J. Fan, Y. Tang, B. Sensale-Rodriguez, C. Yu, and W. Gao, Laser Photonics Rev. **16**, 2200348 (2022).
15. H. Zeng, J. Fan, Y. Zhang, Y. Su, C. Qiu, and W. Gao, Opt. Express **30**, 12712 (2022).
16. G. Qu, G. Cai, X. Sha, Q. Chen, J. Cheng, Y. Zhang, J. Han, Q. Song, and S. Xiao, Laser Photonics Rev. **16**, 2100732 (2022).
17. T. W. Hughes, M. Minkov, V. Liu, Z. Yu, and S. Fan, Appl. Phys. Lett. **119**, 150502 (2021).
18. M. Mansouree, A. McClung, S. Samudrala, and A. Arbabi, ACS Photonics **8**, 455 (2021).
19. J. Skarda, R. Trivedi, L. Su, D. Ahmad-Stein, H. Kwon, S. Han, S. Fan, and J. Vučković, npj Comput. Mater. **8**, 78 (2022).
20. D. Mengu, Y. Zhao, N. T. Yardimci, Y. Rivenson, M. Jarrahi, and A. Ozcan, Nanophotonics **9**, 4207 (2020).
21. K. Makino, K. Kato, Y. Saito, P. Fons, A. V. Kolobov, J. Tominaga, T. Nakano, and M. Nakajima, J. Mater. Chem. C **7**, 8209 (2019).
22. S. Mehrabkhani and T. Schneider, Opt. Express **25**, 30229 (2017).