optica

Three-dimensional waveguide interconnects for scalable integration of photonic neural networks

Johnny Moughames,[1] Xavier Porte,[1,*] Michael Thiel,[2] Gwenn Ulliac,[1] Laurent Larger,[1] Maxime Jacquot,[1] Muamer Kadic,[1] and Daniel Brunner[1]

[1]Institut FEMTO-ST, Université Bourgogne Franche-Comté CNRS UMR 6174, Besançon, France
[2]Nanoscribe GmbH, Hermann-von-Helmholtz-Platz 6, 76344 Eggenstein-Leopoldshafen, Germany
*Corresponding author: javier.porte@femto-st.fr

Received 13 January 2020; revised 20 April 2020; accepted 21 April 2020 (Doc. ID 388205); published 3 June 2020

Photonic waveguides are prime candidates for integrated and parallel photonic interconnects. Such interconnects correspond to large-scale vector matrix products, which are at the heart of neural network computation. However, parallel interconnect circuits realized in two dimensions, for example, by lithography, are strongly limited in size due to disadvantageous scaling. We use three-dimensional (3D) printed photonic waveguides to overcome this limitation. 3D optical couplers with fractal topology efficiently connect large numbers of input and output channels, and we show that the substrate's area and height scale linearly. Going beyond simple couplers, we introduce functional circuits for discrete spatial filters identical to those used in deep convolutional neural networks.

https://doi.org/10.1364/OPTICA.388205

1. INTRODUCTION

The interconnection of numerous input and output channels (IO-channels) is the basic operation behind many applications. A parallel and energy-efficient interconnect has therefore been a desired technology for decades [1,2], finding use in diverse fields such as telecommunication, inter- and intra-chip data buses, and potentially bio-photonics [3]. Most timely, it also is highly desired for connecting layers of deep neural networks (NNs) to efficiently provide the typically large-scale vector matrix products [4].

The integration of such an apparatus is challenging. To achieve parallelism, serial routing is naturally not an option, and a large number of direct physical links connecting the IO-channels is required. Such channel multiplexing can be created in different dimensions, for example, in wavelength or space, and here we address spatial multiplexing. If a direct connection architecture is realized electronically, the capacity of the long connection wires will result in prohibitive energy dissipation and bandwidth limitations [5,6]. There are additional, more practical challenges. Lithographic fabrication typically integrates circuits in two dimensions (2D), and a 2D interconnect's footprint grows quadratic with the number of IO-channels ($N_I|N_O$). Memristive cross-bar or wavelength-division-multiplexing-based approaches illustrate this fundamental relationship.

Optical routing removes the energy dissipation associated with charging the capacity of electronic signaling wires [5], and free-space interconnects with many IO-channels have long been explored [1,2]. Integrated photonic interconnects, however, remain size limited by the unfavorable scaling between area and the number of IO-channels in 2D [7–10]. Such scaling also is found for wavelength-division multiplexing.

We demonstrate the integration of such photonic interconnects in 3D for the first time. Complex 3D-routed waveguides are created by two-photon polymerization [11,12]. We introduce a fractal architecture that efficiently connects many IO-channels, and we demonstrate an integrated photonic interconnect of unreported size hosting 225 input and 529 output channels within a footprint area of only 0.46×0.46 mm^2. As provided by Google in an analysis of its proprietary hardware [13], the neural network model most heavily employed on its servers (MLP0) hosts five layers and 20 million parameters. Assuming a symmetric distribution, this indicates around 2200 neurons in each layer, which sets the future target for N_I and N_O. However, already significantly smaller networks with two layers and 800 neurons achieve 99% recognition rate in the famous MNIST task [14]. We will show that due to linear scaling connecting the layers of such a deep NN in a 3D photonic integrated circuit can be done within a mm^3. Going beyond, we demonstrate a 3D-waveguide architecture implementing nine spatial filters with a Haar convolution kernel [15] of stride and width three. Such convolutional filters represent a fundamental operation of deep convolutional NNs [4]. Our concept is based on mature fabrication technology, which has also been exploited for photonic wirebonding between chips [16,17].

2. SCALING OF INTERCONNECTS

A strategy to overcome many of the bottlenecks currently experienced in NN computation is to realize integrated circuits adhering

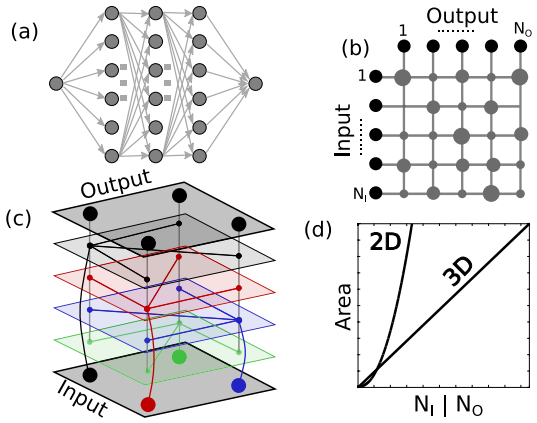

Fig. 1. (a) Topology of a deep neural network. Links between layers of neurons correspond to large-scale interconnects. (b) Crossbar arrays link input and output channels (IO-channels, black dots) in parallel in 2D; IO-channels are arranged along a line. (c) In three dimensions, IO-channels can be arranged in an array, while connections are implemented in the third dimensions. (d) The number of IO-channels of a parallel interconnect scales linearly with size in 3D. In 2D, scalability is significantly worse.

to a NN's complex topology [8,18–22]. As schematically illustrated in Fig. 1(a), a NN is formed by linking large numbers of nonlinear neurons, which often are grouped in layers. It is particularly the parallel intra-neuron interconnect that, despite recent progress [22], still eludes a fully parallel and scalable hardware integration. Most of today's integrated circuits are created via lithography, and are hence restricted mostly to 2D. In cross-bar interconnects [see Fig. 1(b)], routing occurs via punctual contacts between two layers hosting input and output wires. The N_I input and N_O output channels are arranged along a column or row, and hence the overall area scales with $A \propto N_\mathrm{I} N_\mathrm{O}$. This is the general behavior in 2D.

3D additive manufacturing has significantly matured and allows complex structures with nanometric feature sizes [11,23–25]. Crucially, the additional third dimension facilitates simple wiring topologies that are scalable, as schematically illustrated in Fig. 1(c). Input and output ports occupy a dedicated plane each (not rows or columns as in 2D), while the third dimension unlocks a circuit's volume for wiring: for each of the N_I inputs, a dedicated plane hosts all its connections to the N_O outputs. Even in such a simple routing scenario, the system's scaling of area $A \propto N_\mathrm{I}|N_\mathrm{O}$ and height $H \propto N_\mathrm{I}$ becomes linear. The strong impact of 2D versus 3D integration on the scalability of a parallel interconnect is schematically illustrated in Fig. 1(d). Interestingly, the 3D routing strategy has been confirmed by evolution: the most reduced topological property of the human neocortex leverages the same effect. Neurons are located mostly on its surface, while long range connections traverse mostly the volume.

However, 3D routing in electronics is challenging. The required $\propto N_I$ signaling layers make lithographic fabrication prohibitive for the kind of dimensionality demanded by NNs. Heat creation and heat dissipation from such a volumetric circuit's center have additionally been identified as problematic [26]. Disposing of this dissipated energy is a major bottleneck already for the mostly serial von Neumann processors [6], and parallel interconnects for NNs require significantly more such layers and connections. Photonics can overcome this challenge [5,27], which motivates the interest

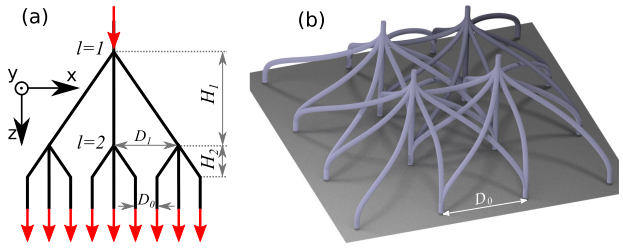

Fig. 2. (a) Design principle of an optical coupler with a fractal geometry. Numerous layers of branching connections can be cascaded, and distances from one layer to the next scale with $\sqrt{b}$, where b is the branching ratio. (b) 3D illustration of a small network hosting simple couplers. Chirality of the connections avoids the intersection of individual waveguides between the input and output ports.

in photonic interconnects [1,2] and ultimately in photonic NNs [8,28–34].

3. 3D INTERCONNECTS OF PHOTONIC WAVEGUIDES

Low-loss 3D printed photonic waveguides have been demonstrated at telecommunication wavelengths [16,17,35,36]. Our waveguides were fabricated using a commercial 3D direct-laser writing system from Nanoscribe GmbH (Photonic Professional GT). A negative tone photoresist "Ip-Dip" was dropped on a fused silica glass substrate ($25 \times 25 \times 0.7$ mm^3) and photopolymerized via two-photon absorption with a $\lambda = 780$ nm femtosecond pulsed laser, focused by a 63X, (1.4 NA) objective lens. After the writing process, the sample was immersed in a PGMEA (1-methoxy-2-propanol acetate) solution for 20 min to remove the unexposed photoresist. Samples were written using the scanning mode based on a goniometric mirror, and the scanning speed on the sample's surface was kept constant at 10 mm/s. As the optimization parameter, we used the writing laser's power. The diameter of individual waveguides was $d \approx 1.2$ μm, and they were spaced by $D_0 = 20$ μm [cf. Fig. 2(a)]. Samples were structurally inspected using a scanning electron microscope (SEM, FEI Quanta 450 W). For optical characterization, we focused a 635 nm laser onto an input waveguide's top surface using a 50X, (NA = 0.8) microscope objective. The mode field diameter of the focused beam was $\simeq 2$ μm, hence larger than the input waveguide's diameter. The emission at the couplers' output ports was collected by a 10X, (NA = 0.30) microscope objective and imaged onto a CMOS camera (iDS U3-3482LE, pixel spacing 2.2 μm) using an achromatic lens with 100 mm focal distance (Thorlabs AC254-100-A), resulting in an optical magnification of 5.6. The details of the optical characterization setup can be found in Supplement 1.

A. Fractal Topology for Fully Connected Layers

Fully or densely connected layers are a principle topology in NNs [4,13]. We adopt a routing strategy based on fractal (self-similar) branching, where each signal "wire" splits into b branches at the $l \in [1, \ldots, L]$ branching layers. Figure 2(a) schematically illustrates such a fractal tree's 2D projection onto the (x, z) plane for $b = 9$ and $L = 2$. An input (top red arrow) is therefore distributed to $N_\mathrm{O} = b^L$ output channels (bottom red arrows), here resulting in $N_\mathrm{O} = 81$. Scaling of N_O is therefore exponential in L, and $N_\mathrm{O} = 6561$ connections are created for each input channel for

$b = 9$ and only $L = 4$ branching layers. The tree's architecture is recursively defined according to the spacing between the N_O output channels $D_L = D_0$ and height $H_L \propto D_0$. The dimensions inside the bifurcation layers $l < L$ are $H_l = \sqrt{b} H_{l+1}$ and $D_l = \sqrt{b} D_{l+1}$. Horizontal and vertical distances scale identically, resulting in constant branching angles throughout the entire circuit.

When arranged in arrays, an additional design step must be applied to avoid crossings from neighboring couplers before layer $l = L$, where waveguides ultimately merge into their respective outputs. Crucially, the translation invariance provided by the scale-free structure aids geometries to avoid such unwanted intersections. These details are illustrated for four neighboring couplers with $b = 9$ and $L = 1$ in Fig. 2(b). We incorporated chirality into the fractal couplers: the b connections from a point in layer l to layer $l + 1$ have a negative curvature in the (x, y) plane, which avoids intersections for connections made in parallel to the x and y axes. Furthermore, avoiding intersections for diagonal links additionally requires curvatures in the z direction. The details of the chiral-fractal topology are given in Supplement 1.

Figure 3(a) shows an SEM image of a 3D fractal coupler array hosting $N_I = 81$ input and $N_O = 121$ outputs, each with $L = 1$ and $b = 9$. We can see that chirality successfully avoids unintended intersections. In Fig. 3(b), we show fractal trees for two bifurcations resulting in 1×81 coupling, with a circuit of $N_I = 9$ inputs and $N_O = 121$ outputs. Noteworthy, the fabrication times for 1×9 and 1×81 fractals are 11 s and 12 min, respectively, and the full structure shown in Fig. 3(b) required approximately 17 min. As for the single-bifurcation-layer 3D coupler, the two-bifurcation-layer couplers are mechanically sound, even though they feature waveguide sections with an aspect ration exceeding 150. This excellent result motivated us to continue and integrate a full-scale interconnect with over 200 inputs, each of which is connected to 81 outputs; see Fig. 3(c).

Figure 4(a) depicts the optical output of a standalone 1×9 fractal coupler, and some of the output ports' optical modes include higher-order Gauss–Laguerre contributions. As our polymer waveguides are freestanding in air, they have a high refractive index contrast of $\Delta n = n_1 - n_0 \approx 0.5$, with $n_1 \approx 1.5$ and $n_0 = 1$, the refractive indices of the polymer and air, respectively. According to the commonly employed approximation $M = 0.5(\pi d \mathrm{NA}/\lambda)^2$, where M is the number of modes supported by the cylindrical waveguide, and $\mathrm{NA}^2 = n_1^2 - n_0^2$, our waveguides support up to $M \approx 22$ optical modes, and the single-mode cut-off wavelength is $\lambda_C = 1755$ nm. However, early stage numerical simulations confirm that mostly the first- and second-order optical modes are excited, which agrees with our experimental results. We would like to point out that the high refractive index contrast allows exceptionally narrow bending radii, which facilitates compact integrated photonic circuits.

We analyzed 1×9 and 1×81 standalone couplers (three each) with respect to the relative power distribution at their output ports, and statistical information is given in Figs. 4(b) and 4(c). For the 1×9 couplers, we find that $(42 \pm 4)\%$ of the total optical output power is provided by the central waveguide, with the remaining ~58% quite evenly distributed among the off-center ports [see Fig. 4(b)]. For 1×81 couplers, $(33 \pm 6)\%$ of light is contained in the central waveguide [Fig. 4(c)]. This is not quite the square of the 1×9 ratio, indicating that cascading our bifurcating waveguides cannot be fully approximated simply by linearly multiplying

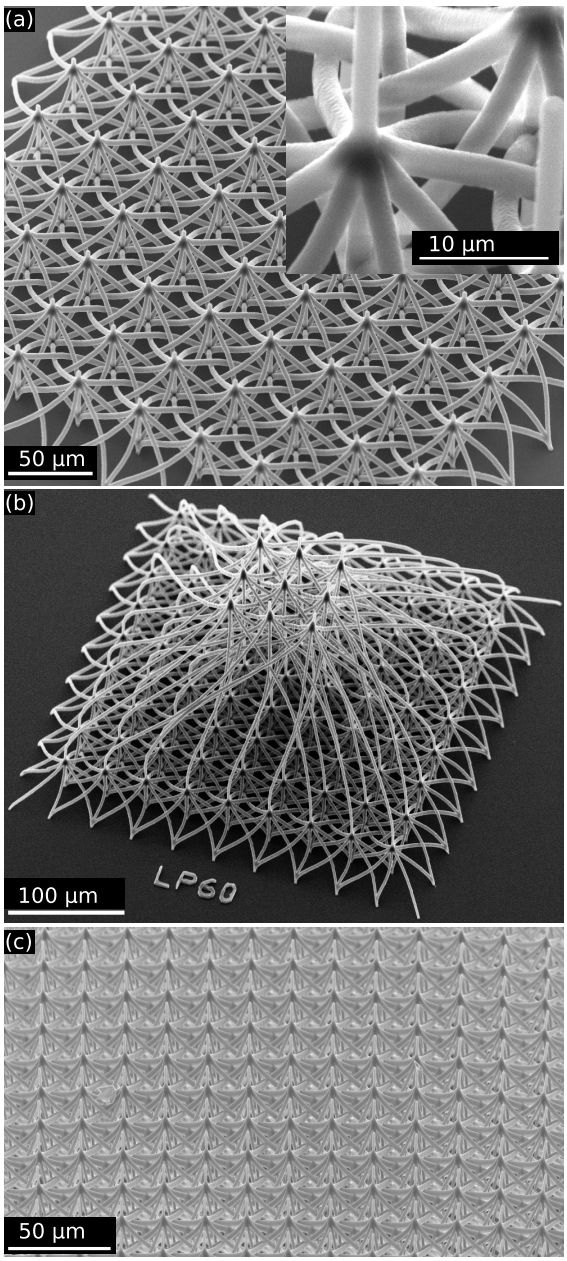

Fig. 3. (a) SEM micrograph (15 kV, 40°) of an array of 1×9 couplers hosting nine elements, including an inset with higher magnification showing the low surface roughness. (b) SEM micrograph (10 kV, 40°) image of nine parallel 1×81 couplers featuring two bifurcation layers. (c) SEM micrograph (10 kV, 40°) of a large-scale array of 1×81 couplers, showing an area containing 15×15 units.

the coupling ratios of the individual components. Higher-order modes therefore appear to have an impact upon the splitting ratios. Overall, the asymmetric splitting ratio is most likely caused by the geometry, and in particular by the branching angles of our waveguide couplers.

Finally, we used the camera images to characterize the optical losses, where the injection spot focused onto the glass-substrate's top surface was used as the calibration reference. The average optical losses for individual 1×9 couplers are 5.5 dB, which rise to 10.6 dB for 1×81 couplers, averaging over 16 printed structures each. These losses comprise optical injection losses I (not induced by the structure), propagation losses P, and losses induced at

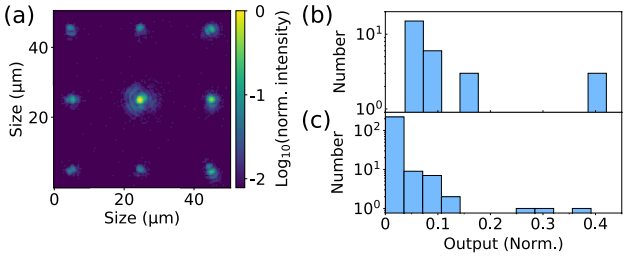

Fig. 4. (a) Optical transmission through a single bifurcation layer 1 × 9 coupler, with intensity color-coded on a logarithmic scale. Histograms of the relative output intensity distribution for the 1 × 9 (b) and 1 × 81 (c) coupler. Statistical information obtained from three couplers each.

each coupling or bifurcation point C. The fractal design principle allows us to determine each of these contributions. As previously discussed, angles of the different bifurcation layers remain constant due to the employed scale-free design principle. This results in identical geometries for all bifurcation points in all topologies, and hence we can assume losses C induced at each waveguide bifurcation point to be identical in all tested structures. Furthermore, we have characterized three 1 × 9 couplers with $\sqrt{b}=3$ times larger height H_L and distance D_L. This leaves us with three loss measurements (in dB): the standard 1 × 9 ($L_{1\times9}=I+C+P$), the three times larger 1 × 9 ($\tilde{L}_{1\times9}=I+C+3P$), and the 1 × 81 ($L_{1\times81}=I+2C+4P$) couplers, and we obtain $I=2.71$ dB, $P=1.14$ dB, and $C=1.67$ dB. The single 1 × 9 and 1 × 81 couplers therefore exhibit 2.8 dB and 7.9 dB losses, respectively, and considering the average propagation length of 51 μm of a single fractal coupler leads to $P\approx20$ dB/mm. Finally, we have also characterized 1 × 9 couplers as part of a large array, as shown in Fig. 3(a). Losses increased from the 2.8 dB for the isolated to 5.4 dB for the coupler embedded in a large array. This might be due to the additional, inverted bifurcating points just at the couplers' outputs.

B. Haar Filters

The previously discussed highly connected couplers are typically required in the final layers of deep NNs. However, their first layers often highlight structural aspects of input information by tailored, local connection topologies. Examples are convolutional NNs commonly employed in object recognition [4]. Prominent convolution kernels are so-called Haar filters. These feature 2D Boolean entries, and this simplification creates a sparse representation of information contained in images, which is a crucial operation for NNs to be able to generalize to unseen test data [15]. We schematically illustrate in- and output properties of nine exemplary Haar filters (F1–F9) in Fig. 5. There, each filter kernel's 3 × 3 Boolean weights (0: dark, 1: bright) are illustrated as input, and each filter is assigned to a single and unique output port.

We developed a 3D routing topology, schematically illustrated on the right in Fig. 5, to realize the nine Haar filters. Even in 3D this is challenging, which can be appreciated from the intricate network of connections, and we have used the same chirality rules discussed for the fractal branching. The realization shown here is according to a geometric arrangement between each filter and its corresponding output waveguide for which the smallest number (six) of intersecting waveguides was found. We then manually modified individual waveguides to remove the remaining intersections. In order to better illustrate the operating principle, we

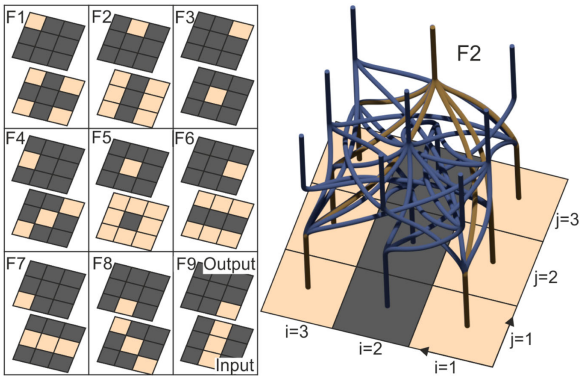

Fig. 5. Schematic illustration of the input–output mapping of nine Haar filters (F1–F9) with Kernel width and stride three. A 3D-printed waveguide architecture realizing all nine filters in parallel is shown on the left, with filter F2 highlighted in orange. The highlighted structure implements filter F2.

have highlighted the connection topology of filter F2 in orange. For each filter, the input ports weighted by one are wired directly to the filter's output. For incoherent injection into 3 × 3 single-mode input waveguides, the intensity at the filter's output should therefore be proportional to the overlap between its Boolean weights and the input.

In Fig. 6(a), we show the SEM image of the 3D printed spatial filtering interconnect realizing nine Haar filters. Waveguides feature smooth surfaces, and the overall structure is stable. However, one can identify a tendency that output waveguides with few connections start leaning outwards. Figure 6(b) shows a densely multiplexed array of Haar filter units. Such an interconnect would implement the convolution of a 21 × 21-pixel input image simultaneously with filters F1–F9. As the individual filter units do not overlap in space, the implemented convolution has stride three.

Figure 6(c) depicts the optical characterization of the filters' connectivity using the same procedure as for the fractal optical couplers. The individual sub-panels correspond to the transmission through a different filter (F1–F9) when injecting light into the output port. The optical characterization was therefore carried out in backward direction. We opted for this procedure since output intensities of individual filters correspond to the filter's Kernel only in the backward direction. In forward direction, one would have to iteratively inject into the individual input ports and then sum the output intensities of the different injections, which is possible in principle yet less systematic. Generally, we find an excellent agreement between the designed filter kernels and the intensities recorded in the reverse propagation direction. The different loss mechanisms obtained for the fractal couplers are consistently reproduced for the Haar filters, with the peculiarity that each coupler exhibits distinct coupling losses C. This, however, is to be expected; different filters rely on specific connection degrees and topologies as well as different branching angles.

There is some crosstalk from the optically injected port onto the image of the output plane. One cause might be the smaller height of the overall 3D circuitry. Light not collected and guided by the injected waveguide illuminates a smaller area on the circuit's output plane, which in turn results in a higher intensity when imaged onto the camera. The outwards-leaning input connections [see Fig. 6(a)] might additionally contribute. The resulting non-orthogonal illumination of the waveguide's tip will most likely

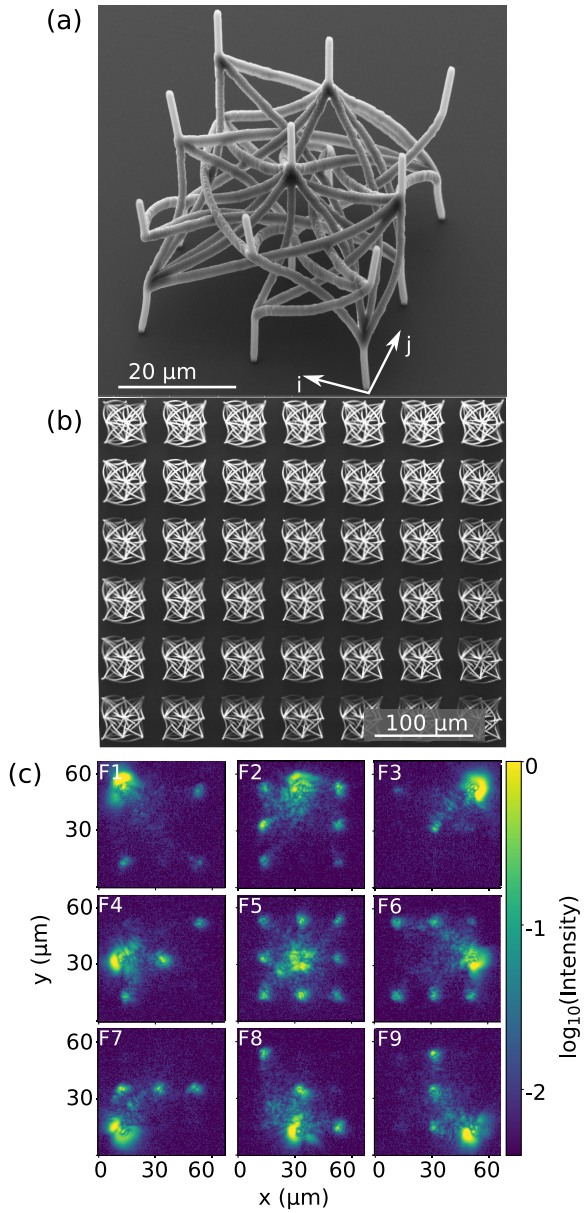

Fig. 6. (a) SEM micrograph (10 kV, 40°) of a single Haar filter. (b) Full micrograph (5 kV, 0°) of a large array hosting spatial filtering for connecting layers of a convolutional neural network. (c) Optical characterization of the filter's connection topology, injection at the output port, and recording the input ports emission.

reduce the injection efficiency and therefore increase the crosstalk of uncollected light to the output plane. For a fully integrated system, this crosstalk will potentially be reduced significantly. Inputs will in most cases be provided by optical fibers or waveguides arriving from an earlier stage of the optical system, for example, when using a fiber bundle for collecting an input image. This will also be true for the filter's output, which will be connected to some fiber or waveguide for further processing downstream. Finally, we have also characterized the filters in the forward direction. However, as in the backward direction, the individual responses are not perfectly symmetric due to the multimode splitters, causing the aggregated responses to not yet correspond well to their target. This highlights the importance of single-mode splitters with more symmetric splitting ratios.

4. DISCUSSION AND CONCLUSION

We successfully demonstrated complex and large-scale 3D photonic interconnects. Waveguides with an approximate diameter of 1.2 µm were created by direct-laser writing based on two-photon polymerization. Using this novel integration strategy, we demonstrated intricate 3D routing topologies for large-scale, highly connected, and convoluting optical interconnects. These exemplary architectures were oriented mostly towards application in NNs, where such interconnects can realize the large-scale vector matrix products fully in parallel with picosecond latency and potentially low energetic cost [10]. It is the first time that such complex and large-scale integrated optical interconnects have been created in 3D. Most interestingly, as mentioned in the introduction, a commercial NN heavily employed by Google features around 2200 neurons in each layer. Using our fractal 3D waveguide methodology, the entire interconnect for linking two layers hosting 2200 neurons would fit inside 1 mm^3! This ignores additional hardware and phase shifters for re-programmability [7], yet it impressively demonstrates the potential of 3D photonic integration. Phase control functionality, and hence reconfigurability of the device, can in the future be added using phase change material coatings [37,38] based on atomic layer deposition [39]. Finally, a NN will require nonlinear elements, and densely spaced arrays of semiconductor lasers are a promising approach [40] with a compatible laser integration density. Future work will have to elaborate on different branching topologies, such as 1×4, for which we expect less asymmetric splitting ratios, and to carefully optimize the couplers' parameters such as the transition's duration when morphing from a single into numerous waveguides. More advanced techniques, such as critical point drying, remove surface tension during the delicate development process, which would allow an even higher integration density.

As our concept scales linearly in size, it allows for novel routing topologies, which in turn will create new opportunities for integrated special purpose NN chips. Here, either complete implementations of NNs or the use of the photonic interconnect purely as a NN accelerator is a possibility [10]. However, there is wider relevance for computing. The end of Moore's and in particular Dennart's scaling is arguably induced by energy penalties of a processor's electronic signaling wires. Photonic routing could prolong the scaling of classical electronic (or now: opto-electronic) von Neumann processors, and these ideas can be expanded to intra- or inter-chip connections.

Future applications can be expected in many other fields, and telecommunication interconnects are certainly a highly interesting possibility [7]. For coherent communication, as in many other contexts, single-mode and polarization preserving propagation is essential. Single-mode guiding at 1550 nm can be achieved either by further reducing the waveguide diameter (to 0.82 µm for the same cut-off wavelength as a SMF28 fiber) or by adding a lower refractive index cladding layer. Modifying a waveguide's cross section is a proven method for maintaining optical polarization. All these advances can be achieved via more elaborate 3D printing protocols or simply by additional deposition techniques, and we believe that no principle mechanisms prohibit this development.

Ultimately, we have demonstrated the first large-scale 3D printed photonic circuit board. The here reported findings are based on the first demonstrations of several, complex 3D photonic circuits, and performance as well as topologies offer significant

potential for further improvements. Beyond losses, it is in particular the asymmetric splitting ratios that deserve further attention, even though such an imbalance can, by a certain degree, be compensated for by using phase-tunable topologies [7] or by the implementation of electro-optical and all-optical phase change layers with a memristive effect [38].

Most importantly, we have addressed the non-scalability of parallel and integrated interconnects for the first time. In order to fully benefit from this new substrate, its functionalization is essential. External control over a waveguide section's phase delay would enable unitary optical transformations on a scalable substrate [7] or by the direct implementation of electro-optical and all-optical phase change layers with a memristive effect. An extension by active or nonlinear photonic elements will establish a new type of photonic device. And finally, small-scale low-bandwidth 3D printed polymer circuits are actively considered in electronics, for example, for wareables [41].

Funding. Conseil régional de Bourgogne-Franche-Comté and Agence Nationale de la Recherche (ANR-17-EURE- 0002); Volkswagen Foundation (NeuroQNet I, NeuroQNet II); H2020 Marie Skłodowska-Curie Actions (713694); Agence Nationale de la Recherche (ANR-15-IDEX-03).

Acknowledgment. The authors thank Marina Raschetti for technical support.

Disclosures. M.T. works at the company Nanoscribe. The other authors declare no conflicts of interest.

See Supplement 1 for supporting content.

REFERENCES

1. J. Shamir, H. J. Caulfield, and R. B. Johnson, "Massive holographic interconnection networks and their limitations," Appl. Opt. **28**, 311 (1989).
2. H. Lee, X. Gu, and D. Psaltis, "Volume holographic interconnections with maximal capacity and minimal cross talk," J. Appl. Phys. **65**, 2191–2194 (1989).
3. D. Choudhury, D. K. McNicholl, A. Repettl, I. Gris-Sanchez, T. A. Birks, Y. Wiaux, and R. R. Thomson, "Compressive optical imaging with a photonic lantern," arXiv:1903.01288 (2019).
4. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature **521**, 436–444 (2015).
5. D. A. Miller, "Attojoule optoelectronics for low-energy information processing and communications," J. Light. Technol. **35**, 346–396 (2017).
6. H. Esmaeilzadeh, E. Blem, R. R. S. Amant, K. Sankaralingam, and D. Burger, "Dark silicon and the end of multicore scaling," IEEE Micro **32**, 122–134 (2012).
7. D. A. B. Miller, "Perfect optics with imperfect components," Optica **2**, 747 (2015).
8. Y. Shen, N. C. Harris, S. Skirlo, M. Prabhu, T. Baehr-Jones, M. Hochberg, X. Sun, S. Zhao, H. Larochelle, D. Englund, and M. Soljacic, "Deep learning with coherent nanophotonic circuits," Nat. Photonics **11**, 441–446 (2017).
9. T. W. Hughes, M. Minkov, Y. Shi, and S. Fan, "Training of photonic neural networks through in situ backpropagation and gradient measurement," Optica **5**, 864 (2018).
10. H. Peng, M. A. Nahmias, T. F. de Lima, A. N. Tait, and B. J. Shastri, "Neuromorphic photonic integrated circuits," IEEE J. Sel. Top. Quantum Electron. **24**, 1–15 (2018).
11. M. Deubel, G. von Freymann, M. Wegener, S. Pereira, K. Busch, and C. M. Soukoulis, "Direct laser writing of three-dimensional photonic-crystal templates for telecommunications," Nat. Mater. **3**, 444–447 (2004).
12. L. Yang, A. Münchinger, M. Kadic, V. Hahn, F. Mayer, E. Blasco, C. Barner-Kowollik, and M. Wegener, "On the Schwarzschild effect in 3D two-photon laser lithography," Adv. Opt. Mater. **7**, 1901040 (2019).
13. N. P. Jouppi, C. Young, N. Patil, D. Patterson, G. Agrawal, R. Bajwa, S. Bates, S. Bhatia, N. Boden, A. Borchers, R. Boyle, P.-L. Cantin, C. Chao, C. Clark, J. Coriell, M. Daley, M. Dau, J. Dean, B. Gelb, T. V. Ghaemmaghami, R. Gottipati, W. Gulland, R. Hagmann, C. R. Ho, D. Hogberg, J. Hu, R. Hundt, D. Hurt, J. Ibarz, A. Jaffey, A. Jaworski, A. Kaplan, H. Khaitan, A. Koch, N. Kumar, S. Lacy, J. Laudon, J. Law, D. Le, C. Leary, Z. Liu, K. Lucke, A. Lundin, G. MacKean, A. Maggiore, M. Mahony, K. Miller, R. Nagarajan, R. Narayanaswami, R. Ni, K. Nix, T. Norrie, M. Omernick, N. Penukonda, A. Phelps, J. Ross, M. Ross, A. Salek, E. Samadiani, C. Severn, G. Sizikov, M. Snelham, J. Souter, D. Steinberg, A. Swing, M. Tan, G. Thorson, B. Tian, H. Toma, E. Tuttle, V. Vasudevan, R. Walter, W. Wang, E. Wilcox, and D. H. Yoon, "In-datacenter performance analysis of a tensor processing unit," arXiv:1704.04760 (2017), pp. 1–17.
14. P. Simard, D. Steinkraus, and J. Platt, "Best practices for convolutional neural networks applied to visual document analysis," in *Proceedings of the Seventh International Conference on Document Analysis and Recognition* (2003), Vol. **343**, pp. 958–963.
15. K. Huang and S. Aviyente, "Sparse representation for signal classification," in *Advances in Neural Information Processing Systems 19*, B. Schölkopf, J. C. Platt, and T. Hoffman, eds. (MIT, 2007), pp. 609–616.
16. N. Lindenmann, G. Balthasar, D. Hillerkuss, R. Schmogrow, M. Jordan, J. Leuthold, W. Freude, and C. Koos, "Photonic wire bonding: a novel concept for chip-scale interconnects," Opt. Express **20**, 17667 (2012).
17. C. Koos, J. Leuthold, W. Freude, N. Lindenmann, S. Koeber, G. Balthasar, J. Hoffmann, T. Hoose, P. Huebner, D. Hillerkuss, and R. Schmogrow, "Photonic wire bonding: connecting nanophotonic circuits across chip boundaries," Proc. SPIE **8613**, 86130W (2013).
18. L. Appeltant, M. C. Soriano, G. V. D. Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, I. Fischer, G. Van der Sande, J. Danckaert, S. Massar, J. Dambre, B. Schrauwen, C. R. Mirasso, and I. Fischer, "Information processing using a single dynamical node as complex system," Nat. Commun. **2**, 468 (2011).
19. M. Nahmias and B. Shastri, "A leaky integrate-and-fire laser neuron for ultrafast cognitive computing," IEEE J. Sel. Top. Quantum Electron. **19**, 1–12 (2013).
20. G. Van der Sande, D. Brunner, and M. C. Soriano, "Advances in photonic reservoir computing," Nanophotonics **6**, 561–576 (2017).
21. D. Brunner, P. Antonik, and X. Porte, "Introduction to novel photonic computing," in *Photonic Reservoir Computing, Optical Recurrent Neural Networks*, B. Daniel, M. C. Soriano, and G. Van der Sande, eds. (De Gruyter, 2019), pp. 1–32.
22. A. Neckar, S. Fok, B. V. Benjamin, T. C. Stewart, N. N. Oza, A. R. Voelker, C. Eliasmith, R. Manohar, and K. Boahen, "Braindrop: a mixed-signal neuromorphic architecture with a dynamical systems-based programming model," Proc. IEEE **107**, 144–164 (2019).
23. J. Moughames, S. Jradi, T. M. Chan, S. Akil, Y. Battie, A. E. Naciri, Z. Herro, S. Guenneau, S. Enoch, L. Joly, J. Cousin, and A. Bruyant, "Wavelength-scale light concentrator made by direct 3D laser writing of polymer metamaterials," Sci. Rep. **6**, 33627 (2016).
24. T. Bückmann, N. Stenger, M. Kadic, J. Kaschke, A. Frölich, T. Kennerknecht, C. Eberl, M. Thiel, and M. Wegener, "Tailored 3D mechanical metamaterials made by dip-in direct-laser-writing optical lithography," Adv. Mater. **24**, 2710–2714 (2012).
25. G. von Freymann, A. Ledermann, M. Thiel, I. Staude, S. Essig, K. Busch, and M. Wegener, "Three-dimensional nanostructures for photonics," Adv. Funct. Mater. **20**, 1038–1052 (2010).
26. V. Venkatadri, B. Sammakia, K. Srihari, and D. Santos, "A review of recent advances in thermal management in three dimensional chip stacks in electronic systems," J. Electron. Packag. **133**, 041011 (2011).
27. A. W. Lohmann, "Principles of optical computing," in *Nonlinear Optics and Optical Computing*, (Springer, 1990), pp. 151–157.
28. N. H. Farhat, D. Psaltis, A. Prata, and E. Paek, "Optical implementation of the Hopfield model," Appl. Opt. **24**, 1469 (1985).

29. L. Larger, M. C. Soriano, D. Brunner, L. Appeltant, J. M. Gutierrez, L. Pesquera, C. R. Mirasso, and I. Fischer, "Photonic information processing beyond Turing: an optoelectronic implementation of reservoir computing," Opt. Express **20**, 3241–3249 (2012).
30. F. Duport, B. Schneider, A. Smerieri, M. Haelterman, and S. Massar, "All-optical reservoir computing," Opt. Express **20**, 22783–22795 (2012).
31. D. Brunner, M. C. Soriano, C. R. Mirasso, and I. Fischer, "Parallel photonic information processing at gigabyte per second data rates using transient states," Nat. Commun. **4**, 1364 (2013).
32. K. Vandoorne, P. Mechet, T. Van Vaerenbergh, M. Fiers, G. Morthier, D. Verstraeten, B. Schrauwen, J. Dambre, and P. Bienstman, "Experimental demonstration of reservoir computing on a silicon photonics chip," Nat. Commun. **5**, 1–6 (2014).
33. D. Pierangeli, G. Marcucci, and C. Conti, "Large-Scale photonic Ising machine by spatial light modulation," Phys. Rev. Lett. **122**, 213902 (2019).
34. E. Khoram, A. Chen, D. Liu, L. Ying, Q. Wang, M. Yuan, and Z. Yu, "Nanophotonic media for artificial neural inference," Photon. Res. **7**, 823 (2019).
35. J. Pyo, J. T. Kim, J. Lee, J. Yoo, and J. H. Je, "3D printed nanophotonic waveguides," Adv. Opt. Mater. **4**, 1190–1195 (2016).
36. A. Nesic, M. Blaicher, T. Hoose, A. Hofmann, M. Lauermann, Y. Kutuvantavida, M. Nöllenburg, S. Randel, W. Freude, and C. Koos, "Photonic-integrated circuits with non-planar topologies realized by 3D-printed waveguide overpasses," Opt. Express **27**, 17402 (2019).
37. C. Rios, M. Stegmaier, P. Hosseini, D. C. Wang, T. C. Scherer, C. Wright, H. Bhaskaran, and W. E. Pernice, "Integrated all-photonic non-volatile multi-level memory," Nat. Photonics **9**, 725–732 (2015).
38. Y. Zhang, J. B. Chou, J. Li, H. Li, Q. Du, A. Yadav, S. Zhou, M. Y. Shalaginov, Z. Fang, H. Zhong, C. Roberts, P. Robinson, B. Bohlin, C. Ríos, H. Lin, M. Kang, T. Gu, J. Warner, V. Liberman, K. Richardson, and J. Hu, "Broadband transparent optical phase change materials for high-performance nonvolatile photonics," Nat. Commun. **10**, 1–9 (2019).
39. A. Frölich and M. Wegener, "Spectroscopic characterization of highly doped ZnO films grown by atomic-layer deposition for three-dimensional infrared metamaterials [Invited]," Opt. Mater. Express **1**, 883 (2011).
40. T. Heuser, J. Große, A. Kaganskiy, D. Brunner, and S. Reitzenstein, "Fabrication of dense diameter-tuned quantum dot micropillar arrays for applications in photonic information processing," J. Appl. Phys. **3**, 1–6 (2018).
41. Y.-G. Park, H. Min, H. Kim, A. Zhexembekova, C. Y. Lee, and J.-U. Park, "Three-dimensional, high-resolution printing of carbon nanotube/liquid metal composites with mechanical and electrical reinforcement," Nano Lett. **19**, 4866–4872 (2019).