

Scalable 3D Silicon Photonic Electronic Integrated Circuits and Their Applications

Yu Zhang , Anirban Samanta , Kuanping Shang , Member, IEEE, and S. J. Ben Yoo , Fellow, IEEE

(Invited Paper)

Abstract—This paper investigates the opportunity and challenges of 3D silicon photonic electronic integrated circuits (3D EPICs) scaling to wafer-scale and beyond. The continuing demand for more data and information is driving new computing, communications, imaging, and information processing at higher throughput and energy-efficiency at lower manufacturing cost. The newly developed 3D silicon photonic devices including vertical U-turns and vertical lightpipes enable through-silicon-optical-vias (TSOVs) that can interconnect multiple layers of 2D silicon photonic electronic integrated circuits consistently with the industry’s through-silicon-via (TSV) based 3D electronic integrated circuit manufacturing. Heterogeneous integration technology based on transfer printing allows wafer-scale (or beyond wafer-scale) heterogeneous integration of dissimilar materials (e.g., III-V semiconductor layers) on silicon at room temperature and 3D ultrafast laser inscription technologies allow arbitrary optical interfaces for high-density input/output between the 3D EPICs and many strands of multi-core-fibers (MCFs). We will discuss applications in neuromorphic computing, 3D LiDAR, photonic-integrated-interferometric-telescopes, 3D-fine-grain-memory, and 3D processor-memory realized by large-scale 3D EPICs.

Index Terms—3D integration, photonic integrated circuit, photonic electronic integration, silicon photonics, nanophotonic fabrication.

I. INTRODUCTION

EVER increasing demands for more information is driving the need for better data processing capabilities with higher speeds, higher energy-efficiency, and greater scalability without significant increases in the cost of manufacturing and ownership. While the Moore’s Law has continued its remarkable scaling over 40 years [4] to correctly predict scaling from hundred transistors per die in pre-1970s to over ten billion transistors per die today, the Dennard’s law [6] which predicted energy-efficiency (performance per watt) to also exponentially improve at roughly

Manuscript received June 17, 2019; revised January 13, 2020; accepted January 31, 2020. Date of publication February 21, 2020; date of current version March 19, 2020. This work was supported in part by Defense Advanced Research Projects Agency (DARPA) under Contract #HR0011-16-C-0106, in part by Office of Naval Research (ONR) under Grant N00014-17-1-2836, in part by the Department of Defense (DoD) under Contract H98230-16-C-0820, and in part by National Science Foundation (NSF) under Grant 1611560. (Corresponding author: Yu Zhang.)

The authors are with the Department of Electronic and Computer Engineering, The University of California, Davis, Davis, CA 95616 USA (e-mail: dyuzhang@ucdavis.edu; asamanta@ucdavis.edu; kshang@ucdavis.edu; sbyoo@ucdavis.edu).

Color versions of one or more of the figures in this article are available online at <https://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTQE.2020.2975656

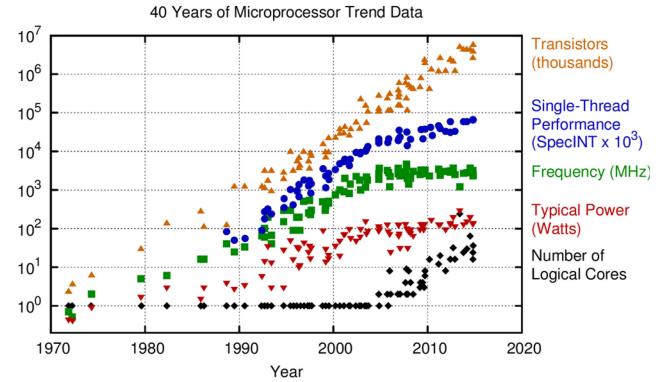


Fig. 1. 40 years of microprocessor trend data showing the continuing trends of exponential growths of transistor integration while power-efficiency, clock-speed, and single-thread performance have ceased its exponential growths in 2005 [1]. On the other hand, new exponential growths in the number of logical cores have emerged around the same time.

the same rate as the Moore’s law have ceased to hold since 2005. As the dimensions of the transistors shrunk to nanoscale, both the leakage current and the threshold voltage variations became non-negligible. Furthermore, the clock speeds could not scale without sacrificing power efficiencies since reduced wire dimensions did not help reduce skin-effects at high frequencies. As Fig. 1 illustrates, the number of transistors on a processor die has continued its exponential growths over the four decades while power-efficiency, clock-speed, and single-thread performance have ceased its exponential growths in 2005 [1]. On the other hand, the new trend of exponential growths in the number of logical cores has emerged around the same time. This new trend is to increase computing performance through involving more transistors and more computing cores running in parallel at lower speeds, rather than to run a few transistors at higher speeds.

Interestingly, as Fig. 2 illustrates, this new trend continued to improve the calculation throughput per \$1000 beyond 2005 despite involving more transistors running at lower clock speeds according to [7]. Now then scaling of the computing throughput, energy-efficiency, and cost will rely strongly on those of the interconnects. Even for computing with a single microprocessor, it was noted that the energy consumed in data movements dominates energy consumption compared to other activities [8]. Hence, massively parallel and high-efficiency interconnects

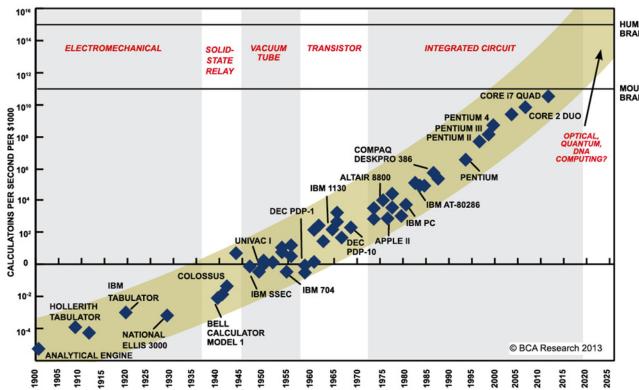


Fig. 2. Estimated computation per second per \$1000 [7].

across multi-core, multi-die, and multi-package are necessary for future information system.

Silicon photonic interconnect is a transformative enabling technology that brings parallelism, high-throughput, power-efficiency, and cost-effectiveness to computing. Integration and packaging between silicon photonics and silicon CMOS components can exploit 2D, 2.5D, or 3D as well as even monolithic integration technologies thanks to the compatibility between the two utilizing the form factors of silicon electronics.

Compared to 2D integration, 2.5D and 3D integration provide higher degrees of functionality for the same footprint. For electronics, power-efficiency and performance also improve with 2.5D and 3D compared to 2D. What about photonics? What technologies are available for 2.5D and 3D photonic integration, and are they compatible with the electronic counterparts? What would 2.5 and 3D electronic-photonics integration look like. What are their benefits and applications? We will investigate these in the following sections.

Section II will discuss 2.5D and 3D electronic integrated circuits, and Section III will cover 2D monolithic electronic-photonics integrated circuits. Section IV will introduce newly developed 3D silicon photonic technologies and Section V will illustrate what 3D photonic and electronic integrated circuits will look like. Section VI will provide four application cases of 3D photonic electronic integrated circuits including neuromorphic computing and imaging. Section VII will summarize this paper.

II. 3D ELECTRONIC INTEGRATED CIRCUITS

Achieving high bandwidth, energy-efficient, and parallel interconnects is difficult on electronic wires because of the electronic impedance, electromagnetic interference induced crosstalk, and skin-effects. In particular, at high speeds, the skin-effect induced ohmic losses and signal distortion increase such that equalizers and repeaters must be included at the expense of additional power consumption and footprints. Especially for memories that often contain long electrical wires 3D integrated memories like high bandwidth memories (HBM)s) or hybrid memory cubes (HMCs) have clear advantages compared to their 2D counterparts. The three main advantages are [9]: (a) lower power consumption: 3D stack allows significant shortening of

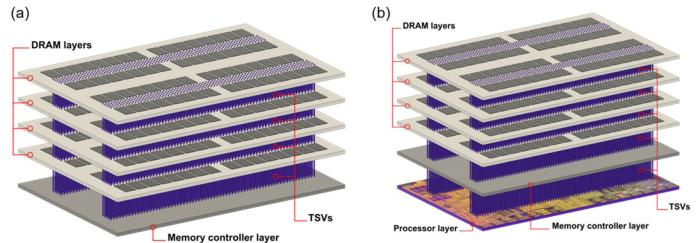


Fig. 3. Schematics of (a) a 3D stacked memory and (b) a 3D stacked processor-memory.

the electrical wires, which results in reduction of resistance and capacitance and in elimination of many repeaters; (b) lower noise: shorter interconnects are less subject to electromagnetic interference and has lower capacitance, lower noise, and lower jitter; (c) higher density: higher packing density in 3D and achieves higher performance per given footprint. One additional advantage is the added communication capability and bandwidth allowed in vertical communications compared to relying on communications in 2D lateral directions alone.

Fig. 3(a) shows a 3D stacked memory example of HBM2 where 8 hi stack of DRAM dies are stacked on a logic layer die. Here, the key enabling technology is through-silicon-vias (TSVs) that allows communications from the logic-layer to the respective layer of the stacked DRAM. (b) includes an additional layer of processor die to facilitate high-bandwidth communications between the processor and the memory, thus relieving the von Neumann bottleneck commonly seen in nearly all computing systems with electronic interconnects in 2D.

The obvious challenges of 3D Electronic ICs are yield and heat density. In terms of yield, the advances in the TSV and manufacturing technologies have enabled 128 layers of fabrication in 2019 and are projected to have 512 layers in 2022 [10].

III. 2D SILICON PHOTONIC-ELECTRONIC INTEGRATED CIRCUITS

Fabrication processes of silicon photonics stems mostly from those of silicon CMOS electronics with a few additional steps such as germanium epitaxial growths for photodetectors, hence monolithic integration of silicon photonics and silicon CMOS electronics can be accomplished by modifying silicon CMOS fabrication processes by adding a number of silicon photonic specific processing steps (e.g., Ge photodetectors and modulators) while combining many common steps (e.g., ion implantation, dry etching, CMP, and metal deposition steps). In fact, such monolithic 2D photonic-electronic integrated circuit fabrication processes have been successfully demonstrated and commercialized.

Fig. 4 is a photograph of monolithic silicon photonic CMOS with integrated Mach-Zehnder modulators, and electronic IC drivers [11] developed by IBM using their CMOS9WG process (~90 nm) in 2015. The monolithic CMOS integrated nanophotonic transmitter obtained a link sensitivity comparable to a 25 Gb/s commercial reference transmitter, exhibiting a 5.2 dB

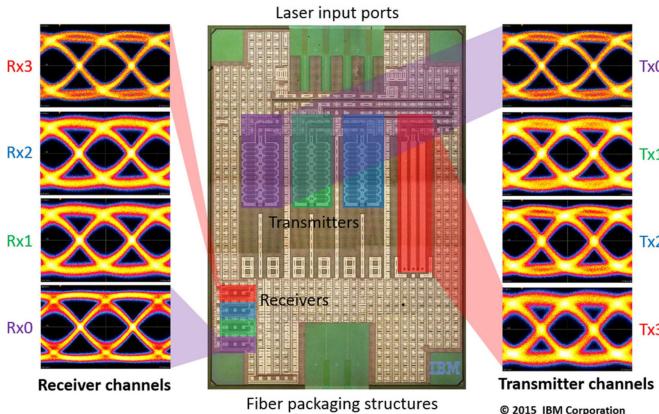


Fig. 4. Photograph of a monolithic integrated silicon photonic CMOS transceiver from IBM [11].

extinction ratio, 4.9 dB insertion loss, and error free operation up to 32 Gb/s [11]. This process has been transferred to GlobalFoundries as a commercially available 90WG process whose PDK now includes transceivers and receivers.

On the other hand, the UC Berkeley-MIT group has successfully utilized the 45 nm SOI BI-CMOS electronic IC process with ‘zero-change’ to incorporate ring modulators and photodetectors in the CMOS to realize similar monolithic silicon photonic electronic ICs [12].

Interestingly, the concept of monolithic 2D photonic-electronic integrated circuit dates back to 1980s when Optoelectronic Integrated Circuits (OEIC) efforts concentrated on utilizing III-V heterojunction bipolar transistor (HBT) platforms to include photodetectors in the reverse-biased base-collector regions and light sources (LEDs and lasers) in the forward-biased emitter-base regions. One advantage the III-V OEICs had compared to the monolithic silicon photonic CMOS integrated circuits is that the OEICs had optical sources from the direct bandgap and electro-optical modulation based on the Pockel’s effect available from the III-V semiconductors. However, unlike silicon that benefits from high-density insulation and passivation material of silicon dioxide through thermal oxidation of silicon itself, III-V materials do not have natural passivation from III-V materials themselves, and must rely on materials based on epitaxial regrowths or ion implantation. The yield of III-V OEIC was far inferior to that of monolithic silicon CMOS photonics due to a number of reasons including challenges in achieving high-quality passivation, material uniformity, process uniformity, and large-scale wafer scale fabrication.

IV. 3D SILICON PHOTONICS

Like the 3D integrated electronics, 3D photonics overcome many limitations of 2D photonics where all photonic components must reside on the same plane. 3D photonic integration on silicon photonic platform plays a significant role because it combines low power consumption, high-density functionalities, and high-yield manufacturing on a CMOS platform already

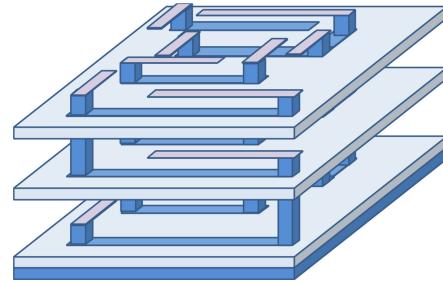


Fig. 5. A conceptual diagram of a 3D silicon photonic integrated circuit analogous to the 3D silicon electronic integrated circuit of Fig. 3.

developed for 3D electronic integration and its natural convergence with the current 2.5D/3D packaging techniques on silicon interposers.

Multilayer integration via evanescent couplers has been demonstrated [13]–[15] and become available at foundry services [16]. This approach not only decreases the overall footprints by the factor of the layer number but also facilitates waveguide routing for large scale systems with multiple inputs and outputs ports.

However, the evanescent couplers for vertical light transportation are typically a few hundred microns long, prohibit the scaling towards future dense integrated low power consumption systems. As the very fine pitch TSV 3D electronic integration are moving towards sub-10 μm pitch [17], [18], it is necessary to develop a corresponding through-silicon-optical-via (TSOV) that can be as small as a typical waveguide size as an enabling technology. With such TSOV interconnecting silicon photonic layers vertically, analogously to the TSVs in the current 3D silicon electronic ICs described in, a 3D silicon photonic integrated circuit as shown in Fig. 5 would be possible. In the following, we will review our recently developed sub- μm^3 volume vertical 3D U-turn TSOV structure utilizing standard silicon photonic fabrication tools [19].

Fig. 6(a) and (b) reveal the detailed structure of the vertical U-shaped coupler as a TSOV. It consists of two 45° reflectors connecting to different silicon waveguiding layers and an inter-layer vertical optical via. The silicon layer thickness is chosen to be 500 nm to be compatible with hybrid integration with other materials like InP for efficient electro-optical modulation [20] and the waveguides are fully-etched to reduce optical coupling to the surrounding elements. Fig. 6(c) shows a finite-difference time-domain (FDTD) simulated light transmission from the bottom silicon layer to the top. With properly set design parameters, the U-shaped coupler could achieve 74% light transmission within a 1 μm^3 volume.

The fabrication of such TSOV involves multiple steps of selectively wet etching, planarization and bonding of another crystalline silicon layer. Fig. 7 details the fabrication process of the vertical U-shaped coupler [19]. Fig. 8(a) shows the perspective view of fabricated vertical U-shaped couplers after partly removing top oxide cladding.

The vertical U-shaped coupler transmission loss is measured using cascaded interlayer coupling test structures. Our current

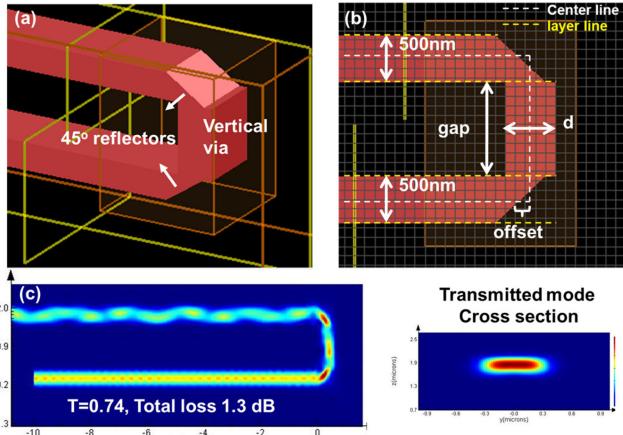


Fig. 6. (a) Perspective and (b) cross-sectional schematic of the ultra-compact vertical U-shaped coupler consisting of a silicon photonic vertical via and two 45° reflectors. (c) FDTD simulated transmission of the vertical U-shaped coupler from the bottom layer to the top layer. Reproduced from [19] © 2018 IEEE.

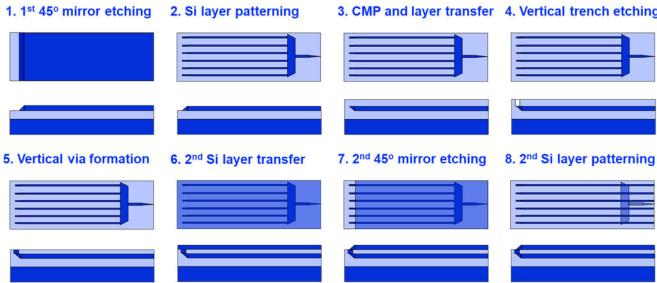


Fig. 7. Fabrication process flow of U-shaped vertical coupler. Reproduced from [19] © 2018 IEEE.

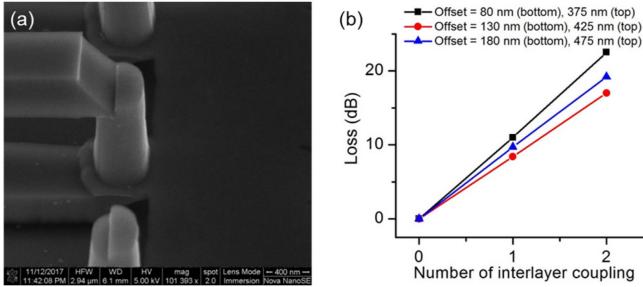


Fig. 8. (a) 15° tiled view SEM pictures of fabricated U-shaped coupler structure with 2 μm pitch. (b) Measured transmission loss of cascaded vertical U-shaped coupler test structures. Reproduced from [19] © 2018 IEEE.

measurement results reveal a minimum loss of 8.4 dB (Fig. 8(b)) per coupler with offset = 130 nm (bottom) and 425 nm (top). We attribute the relatively high loss mostly to a large misalignment between the top 45° reflectors and the vertical vias (~ 300 nm larger than the designed value), and the excessive etching forming an α -Si scatter near the reflector. Our simulation results suggest that these fabrication errors contributed a 4~5 dB additional loss. Using a Vernier ruler based structure for misalignment correction and further controlling the 45° etching undercut using a test structure, we are able to reduce the misalignment to below

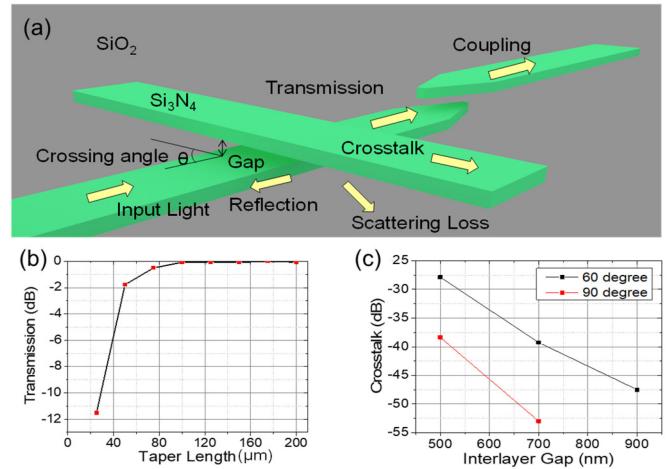


Fig. 9. (a) The schematic view of a waveguide crossing on the bi-layer symmetric Si_3N_4 multilayer platform. Measured (b) transition loss and (c) interlayer crosstalk. Reproduced from [13] © 2015 OSA.

50 nm. We are also improving our annealing process to reduce the α -Si absorption loss and new designs are carrying on to suppress the higher order mode induced loss. With all these methods, we are targeting to achieve low loss coupling towards ~ 1.3 dB as simulated.

Compared with the 2D planar PIC, 3D PIC fabricated with the TSOV has higher cost per area due to the additional four layers of fabrication (vertical via, two 45° mirror etching and 2nd silicon layer). Extra cost includes 4 photomasks, 1 wafer bonding process, multiple deposition and CMP, multiple lithography and etching process and projected lower yield. For certain applications (See Section IV), demonstrated 3D TSOV technology could save at least half of the total area. Therefore, the cost per PIC could be lower for the 3D TSOV PIC.

V. 3D PHOTONIC INTEGRATED CIRCUITS

With the growing needs and development of 3D photonics, large-scale 3D photonic integrated circuits (3D PICs) are being fabricated to achieve complex functionality with low size, weight and power. In this section, we will show several large-scale developed 3D PICs via different techniques.

As we briefly discussed in Section IV, multilayer integration via evanescent couplers has attracted great interests according to the footprint and waveguide routing complexity reduction. Deposited thin film SiN offers a low-cost and multi-purpose solution to this multilayer platform due to its advantage of low-loss and relatively high-index-contrast [13], [21], [22].

For a practical multilayer platform, two essential parameters are affecting the system performance: (1) transition loss from layer-to-layer, (2) optical coupling crosstalk between different layers. Fig. 9(a) shows a typical symmetric two-layer SiN platform. With optimized design, it is possible to achieve low transition loss (<0.1 dB/transition, Fig. 9(b)) and low crosstalk (<-50 dB/crossing, Fig. 9(c)).

We built a bi-layer 27 × 64 star couplers on this platform, Fig. 10(a) shows the image of the fabricated multilayer star

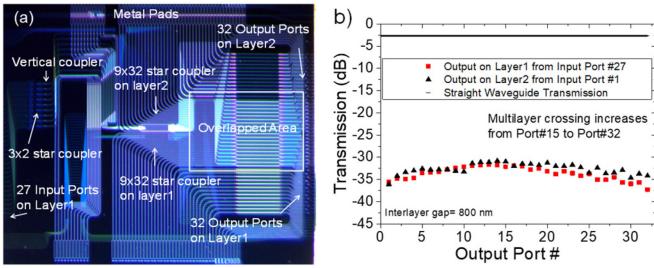


Fig. 10. (a) The schematic view of a waveguide crossing on the bi-layer symmetric Si_3N_4 multilayer platform. Measured (b) transition loss and (c) interlayer crosstalk. Reproduced from [13] © 2015 OSA.

couplers. By employing the multilayer overlapping, the footprint of this device decreases to 40 mm^2 , and no significant performance variation between the top and bottom layer is observed (Fig. 10(b)).

This multilayer platform could also merge devices optimized on different platforms into a single integrated platform [14]. For example, by varying the Si_3N_4 thickness from 50 nm and 200 nm, the optical confinement factor can be greatly adjusted for ultra-low-loss [23] or compact size [24]. Vertically functional devices like arbitrary power splitter in tri-layer could also be demonstrated on this platform [14].

VI. 3D Si-EPIC EXAMPLES AND APPLICATIONS

While the 2D planar PIC dominates the current demonstrated device and system, 3D photonic integration could potentially become attractive and necessary in many very large scale electronic and photonic systems. In particular, we projected that the 3D integration will be important to optical phased array based high performance LIDAR, due to its high fill factor and ease of modularization. For aeronautics and space applications, 3D integration is also attractive for its reduced size and weight. In many high capacity and lower power consumption computing area, 3D integrated optics could offer significant power reduction when replacing the conventional TSV and computing methods.

In this section, we provide four application cases of on-going and potential 3D photonic electronic integrated circuits including neuromorphic computing and imaging.

A. High-Resolution and Ghost-Free LIDAR

Using the waveguide scale TSOV, it is possible to develop an optical phased array (OPA) with element pitch down to $1 \mu\text{m}$ [25] or below. Fig. 11(a) shows a proof-of-concept OPA device with $2 \mu\text{m}$ grating waveguide pitch fabricated by the similar fabrication process shown in Fig. 7. Fig. 11(b)–(e) shows the zoom-in of different session of the 120 channel OPA device (1×128 MMI tree based splitter (Fig. 11(b))), thermal-optical phase modulators (Fig. 11(c)), light transport from top splitter layer to the bottom grating emitter array through the vertical U-shaped coupler (Fig. 11(d))). Current OPA emitting area fill-factor is limited by the thermal crosstalk between the thermo-optic phase modulators. With the recently demonstrated dense phase modulator array, it is possible to realize large fill factor ($\sim 95\%$).

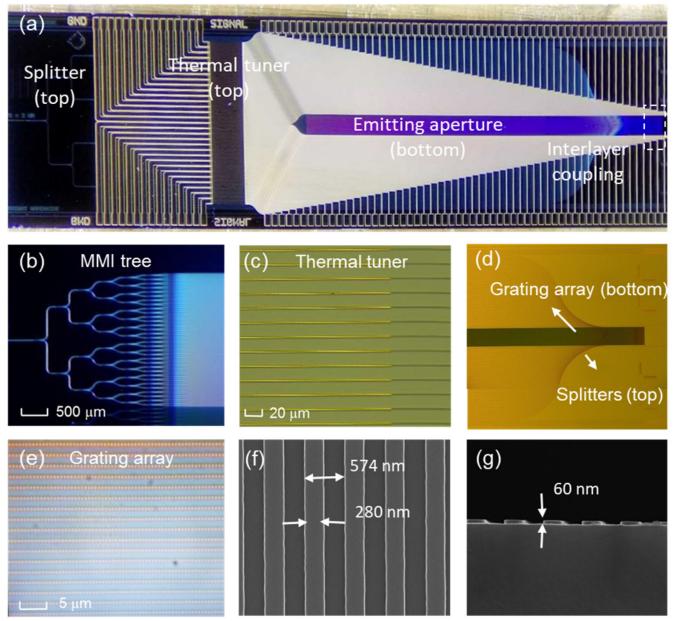


Fig. 11. (a) Optical microscope image of fabricated proof-of-concept 3D integrated silicon photonic unit cell. Zoom-in of (b) the MMI tree based splitter, (c) the heater based thermal tuners, (d) the two silicon photonic layers with U-shaped couplers and (e) the $2 \mu\text{m}$ pitch waveguide grating array. (f) Top view and (g) cross-sectional view SEM pictures of fabricated partially etched SiN assisted grating. Reprinted from [19] © 2018 IEEE.

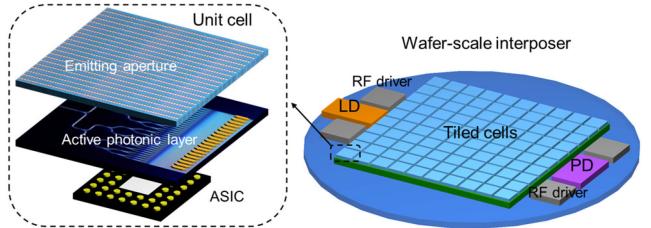


Fig. 11(f) and (g) show the SEM picture of the partially etched SiN assisted weak emitting grating. We measured a grating emitting strength of $\sim 23 \text{ dB/cm}$, corresponds to $>90\%$ optical power radiates out over 5 mm length. It can be further optimized using a chirped design to achieve sharp instantaneous field-of-view [26].

We can treat the demonstrated OPA device as a unit cell and tile it up on a wafer-scale interposer to form a large coherent aperture LIDAR system that can be high-resolution and ghost-free. Fig. 12 illustrates the wafer-scale integration scheme of the LIDAR system. Each unit cell will have two photonic layers, namely emitting aperture on the top surface and all active photonic underneath to be vertically integrated with control ASIC unit that lies in between interposer and the unit cell. A light signal will be distributed to each unit cell through the interposer and transmitting out via the top emitting aperture, and the reflected signals will be collected and relayed to the photodetectors on

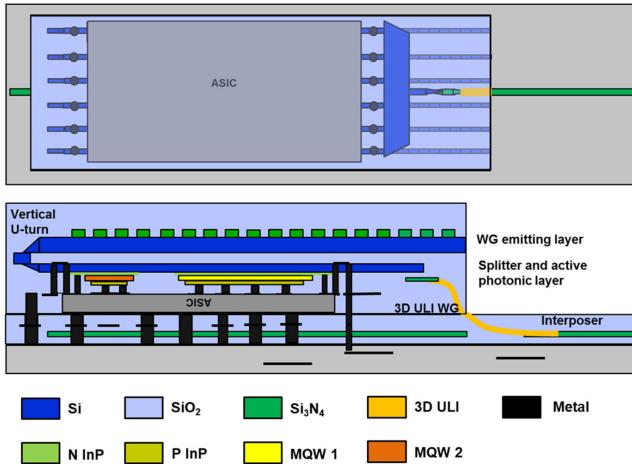


Fig. 13. Schematic of a 3D silicon photonic unit cell on an optical interposer. Reprinted from [19] © 2018 IEEE.

the interposer. Fig. 13 shows the cross-sectional schematic of the proposed 3D EPIC on an optical interposer.

B. 3D EPIC Memory and 3D EPIC-Processor-Memory

Another important application for 3D EPIC is 3D optically connected memory for integrated heterogeneous computing systems. Replacing most of the conventional electrical TSV in the 3D stacked memory (Fig. 3) with the developed TSOV in section IV could not only greatly reduce the footprint occupied by these interconnecting structures, but also increase the communication data capacity if using the wavelength-division-multiplexing (WDM) technology. Fig. 14 shows our proposed 3D stacked of silicon-photonic fine-grain electronic DRAM dies. With the silicon photonic TSOVs connecting each stacked memory layer to the bottom logic layer and deep integrated silicon photonic components on the memory layer, we expect to reduce latency by $2\times$ and energy consumption by $2\times$ compared to the electronic DRAM counterparts.

To fabricate the proposed TSV and TSOV structures, two deep etching through the chip and filling with metal and α -silicon/polysilicon respectively are needed. While the small diameter TSV are well-developed [17], etching a few tens of μm with sub- μm feature remains difficult. $>20:1$ aspect ratio key hole free polysilicon filling has been demonstrated with designed sacrificial layers [27]. Further process development is still needed to maintain single mode operation in the vertical Si via regions to avoid large excess loss.

C. Large-Scale Nanophotonic Neuromorphic Computing

Recently, we presented the architecture, system design, modeling, and simulation results of spiking nanophotonic neural networks capable of achieving hand-writing recognition [3], [5].

As the first task, we designed energy-efficient bio-inspired nanophotonic neurons based on nanophotonic detectors and a nanophotonic laser integrated with nanoelectronics as illustrated in Fig. 15. This simple optoelectronic neuron circuit represents a basic function of neuron action. By replacing the input/output

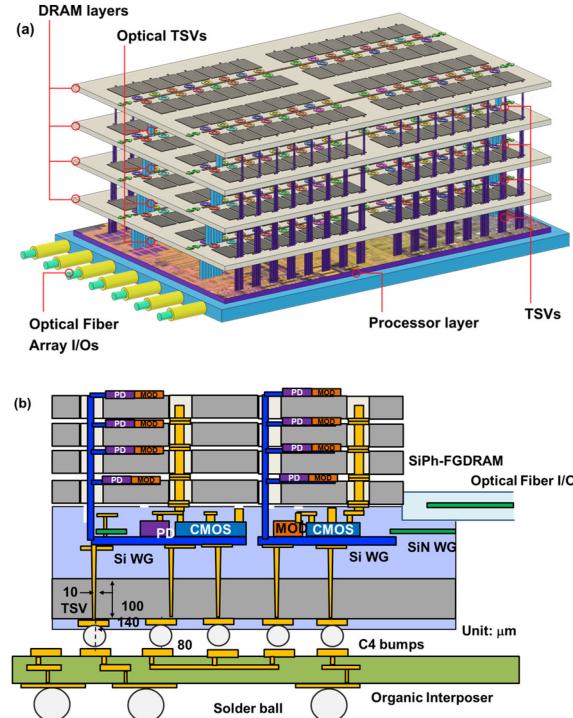


Fig. 14. (a) 3D stacking of silicon-photonic fine-grain electronic DRAM dies with the silicon photonic logic layer at bottom vertically interconnected by through-silicon-optical-vias (TSOVs). (b) Proposed SiPh-FGDRAM utilizing silicon-photonic CMOS DRAM mini-bank dies in multi-stacks including the silicon photonic logic layer at the bottom.

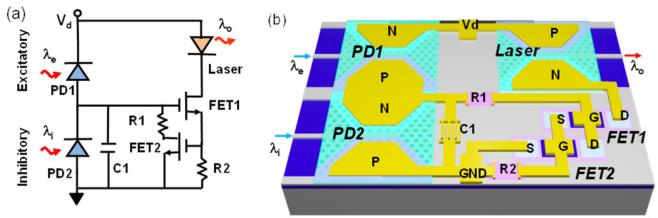


Fig. 15. (a) An example circuit designed for generating bio-inspired spike signals [2], (b) [artwork courtesy of Weidong Zhou] A proposed nanophotonic neuron with the spiking electronic circuit, nanophotonic detectors (PD1 for excitatory and PD2 for inhibitory), a nanophotonic laser, and other elements such as resistors and capacitance (C1 inclusive of the capacitance of PDs and FETs) can be tuned to achieve desired sigmoid response function including temporal and rate coding [3], [5].

interfaces with nanophotonic PDs and lasers, the neuron power consumption can be drastically reduced due to the absence of capacitive charge associated with the interconnect wires [28]. By design of the nanophotonic and FET structures, the static power consumption is extremely low—we expect ~ 2 nW for ~ 2 V supply and the reserve leakage current of ~ 1 nA. By utilizing commercial monolithic CMOS silicon photonic processes (e.g., GlobalFoundries' 90WG and 45CLO) and transfer-printing [29] of quantum-dot (QD) nanolaser structures, our model shows ~ 10 fJ/spike neural links overcoming ~ 19 dB link loss. Utilizing the Izhikevich's model [30], we have encoded and generated photonic spikes of various neurons and input them into the

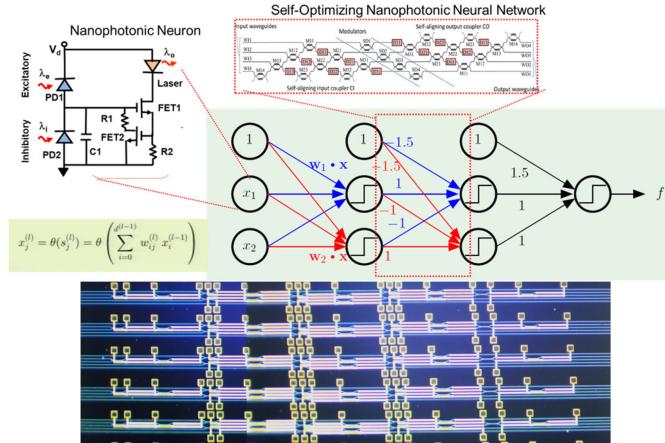


Fig. 16. A single wavelength example of proposed Nanophotonic Neural Network (3-layer single wavelength example) with Nanophotonic Neurons (inset diagram on the left) at each node, and the self-optimizing nanophotonic neural network with 2×2 Mach-Zehnder interferometer (MZI) between each layer. The photograph at the bottom is the fabricated silicon photonic 4×4 nanophotonic synaptic interconnect neural network integrated circuit.

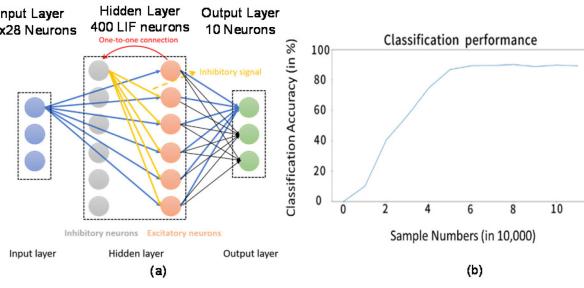


Fig. 17. (a) Handwriting recognition of 28×28 pixel NMIST data with 400 excitatory and 400 inhibitory LIF neurons in the hidden layer and 10 output layer neurons. (b) Classification accuracy vs. training sample number showing $\sim 90\%$ accuracy is possible at $\sim 50,000$ NMIST training data.

LIF optoelectronic neuron of Fig. 15. By creating compact models implemented in Verilog-A, we simulated outputs from the nanoscale optoelectronic (nanophotonic) neuron. Then, as shown in Fig. 16 we constructed a three-layer nanophotonic neural network interconnecting the nanophotonic neuron to simulate the output and investigated hand-writing recognition application.

Fig. 17(a) shows the three layer Spiking Nanophotonic Neural Network (SNNN) including excitatory and inhibitory neurons as illustrated in Fig. 15, and we utilized the BRIAN simulator [31] to conduct the handwriting recognition of 28×28 pixel images. The time-encoded spikes representing the 28×28 pixel hand-writing-images are sent to the 20×20 input nanophotonic neurons, and propagated through the three layer SNNN with the lateral inhibition shown in Fig. 16. In a well-trained network (after ~ 15 iterations) the classification accuracy was about 90%, with more neurons and larger training datasets, higher accuracy can be obtained, while achieving $100\times$ improvements in energy-efficiency and throughput compared to a state-of-the-art electronic neuromorphic counterpart (in this case IBM TrueNorth). At the same time, the 2D nanophotonic neural network is orders

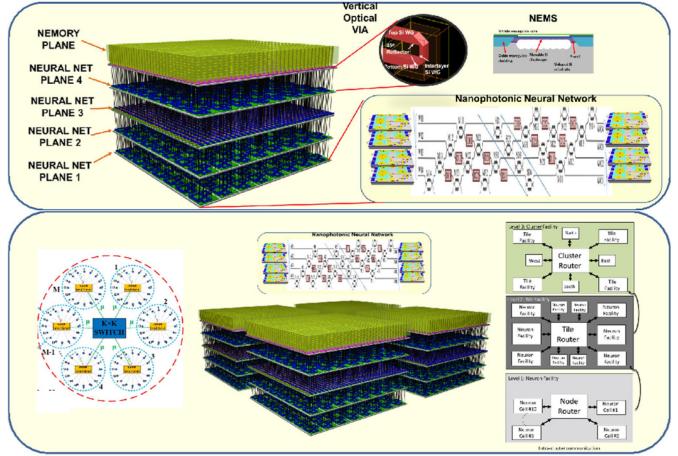


Fig. 18. (Top) 3D integrated neuromorphic system utilizing the TSOV and vertical U-turns interconnecting the multiple 2D nanophotonic neural network planes. (Bottom) multi-3D chip neural networks.

of magnitude larger in footprint area compared to the electronic counterpart. Hence the 3D EPIC realization of the nanophotonic neural network as illustrated in Fig. 18 utilizing TSOVs and TSVs to interconnect multiple 2D neural network planes can be extremely desirable for large-scale nanophotonic neural network systems.

D. Large-Scale Hyperspectral Imaging

The recent demonstration of photonic-integrated-interferometric-telescopes (PIITs) [32], [33] on silicon photonic dies inspire a very large-scale hyperspectral imaging system on a full 300 mm or 450 mm wafer scale or even on a multi-wafer scale. To facilitate this, it is desirable to utilize a hyperspectral imager with vertical apertures. Unfortunately, all photonic integrated spectrometers with high resolution such as arrayed waveguide gratings or Echelle gratings are in the same plane as the wafers. Here, the vertical apertures created in the fabrication steps shown in Fig. 7 and Fig. 8 can create vertical apertures for this purpose. Fig. 19 shows high resolution arrayed waveguide grating (AWG) spectrometer array with vertical apertures and readout circuits and meta lenses integrated in 3D.

VII. HETEROGENEOUS 3D EPIC FABRICATION TECHNIQUES

A. 3D Ultrafast Laser Inscription

3D silicon EPICs will involve very high-density optical interconnects inside the 3D EPIC dies, and their input/output will involve many wavelength channels on many waveguides. Traditional fiber interfaces will unlikely support such high-density I/O, and multi-core-fiber (MCF) array interfaces may be desirable. To interface such high port-count 3D silicon EPIC with a large number of MCFs, we have been investigating 3D ultrafast laser inscription technologies for realizing arbitrary interfaces between the two different systems. Fig. 20 shows the developed 3D ULI producing 16 core fiber interfaces shown in (a1-a2) and 256 core fiber interfaces shown in (b1-b2). On the

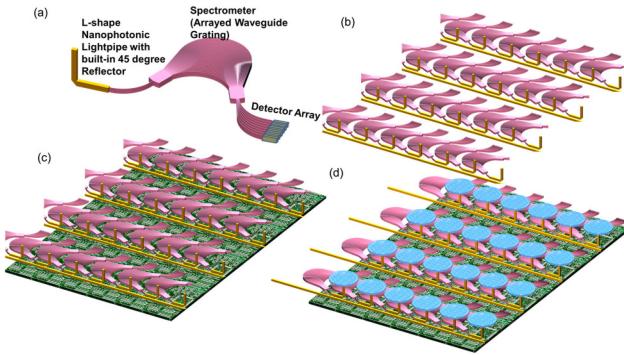


Fig. 19. (a) high-resolution arrayed waveguide grating spectrometer with integrated L-shape nanophotonic lightpipe with integrated 45 degree mirror and integrated photodetector array with possible preamplifiers, (b) a 2 dimensional array of such spectrometers with integrated lightpipes and photodetectors, (c) a 2 D detector array with integrated readout circuits, and (d) integrated hyperspectral imager with integrated metalens array.

opposite end of the 3D ULI block can utilize adiabatic coupling between negative-tapered silicon nitride waveguides (150 or 200 nm thickness) and the ULI waveguide as shown in (c1, c2, and c3) where nearly perfect mode transformation from (c2) the entrance from SiN waveguide to (c3) the exit to 3D waveguide can be possible to achieve 0.04 dB loss. (d) the silicon photonic 3D fan-out chip including the 3D ULI MCF interfaces. Shown is an 8-core MCF with 4 strands producing 32 silicon nitride waveguide interfaces to silicon photonic dies.

B. Transfer-Printing

As silicon is inefficient for light emission and detection in the conventional Datacom/Telecom wavelengths, other materials like InP, GaAs and Ge are needed to achieve full functionality. Although wafer bonding [34] and epitaxial growth on silicon [35] have been developed, these techniques are typically associated with high temperature [35] and might not able be applied to wafer scale due to the wafer size mismatch [34]. With multiple dies-to-wafer technology [36] and CMOS-compatible metallization [37], high performance lasers has been developed in 200 mm silicon photonic platform in foundry with >80% yield.

Micro transfer printing [38], which transfers selectively removed material/devices in a massive parallel manner, enables a new way for such wafer scale integration on many different substrates. Different devices are firstly fabricated on source wafers and underetched to leave the devices anchored to the substrate with tether structures. Devices to be printed are then transferred on the target wafer using an elastomeric stamp. The pick up and printing process are controlled by the contact speed between the stamp and wafers [39].

Being in the micron scale, such transfer overcomes many difficulties that will be encountered for full wafer and offers dense integration with high yield (>99.9%), high throughput (<10 mins to print a 6 inch wafer), relatively accurate alignment ($\pm 1.5 \mu\text{m}$, 3σ) [40] and can be done at room temperature. Based

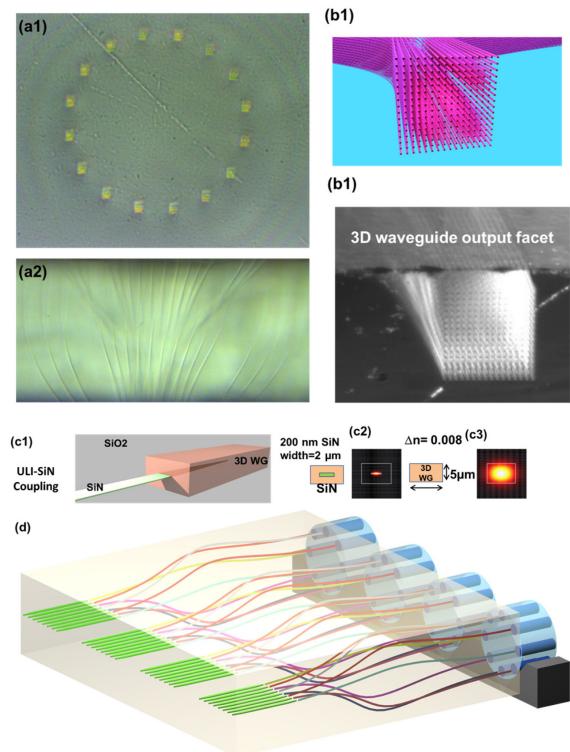


Fig. 20. 3D ULI developed at UC Davis producing 16 core fiber interfaces shown in (a1–a2) and 256 core fiber interfaces shown in (b1–b2). On the opposite end of the 3D ULI block can utilize adiabatic coupling between negative-tapered silicon nitride waveguides (150 or 200 nm thickness) and the ULI waveguide as shown in (c1, c2, and c3) where nearly perfect mode transformation from (c2) the entrance from SiN waveguide to (c3) the exit to 3D waveguide can be possible to achieve 0.04 dB loss. (d) The silicon photonic 3D fan-out chip including the 3D ULI Multi-Core-Fiber (MCF) interfaces. Shown is an 8-core MCF with 4 strands producing 32 silicon nitride waveguide interface to silicon photonic dies.

on this technique, various types of heterogenous integrated photonic devices have been demonstrated on 200 mm SOI platform, including hybrid laser [41], [42], high speed Ge photodetectors [43] and passive filters [44].

VIII. FUTURE DIRECTIONS AND SUMMARY

The rapid increases in data communications, computing, imaging, and information processing continue to drive advanced electronic and photonic integration in 2D and 3D. In particular, we expect to see strong advances in 3D co-integration of photonics and electronics in silicon exploiting the new TSOV capabilities co-existing with TSVs. Heterogeneous integration techniques involving transfer-printing is attractive for scaling to wafer-scale and beyond at moderate processing temperatures and 3D ULI technologies offer arbitrary waveguide interfaces for wafer-scale interposers and high-port-count MCF interfaces.

ACKNOWLEDGMENT

The authors would like to thank all the pioneering works and technical support from the Next Generation Networking

Systems Laboratory members, including but not limited to: B. Guan, R. Scott, Y.-chun Ling, T. Su, G. Liu, K. Zhang, J. Hu, R. Proietti, and P. Fotouhi, and also Profs. L. Coldren, J. Bowers, and J. Klamkin from the University of California, Santa Barbara, Dr. P. Suni and J. R. Colosimo from Lockheed Martin, Prof. S. Palermo from Texas A&M University and Prof. V. Akella and J. Lowe-Power from the University of California, Davis for helpful discussions on computing architecture topics. Fabrication of the devices presented here took place at the Center for Nano-Micro Manufacturing (Davis, CA) and at the Marvell Nanolab Facility (Berkeley, CA). The views, opinions and findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.

REFERENCES

- [1] K. Rupp, “40 years of microprocessor trend data,” 2015. [Online]. Available: <https://www.karlrupp.net/2015/06/40-years-of-microprocessor-trend-data/>
- [2] A. Van Schaik, “Building blocks for electronic spiking neural networks,” *Neural Netw.*, vol. 14, no. 6–7, pp. 617–628, 2001.
- [3] S. J. B. Yoo and D. A. B. Miller, “Nanophotonic computing: Scalable and energy-efficient computing with attojoule nanophotonics,” in *Proc. IEEE Photon. Soc. Summer Topical Meeting Series*, 2017, pp. 1–2.
- [4] G. E. Moore, “Cramming more components onto integrated circuits,” *Proc. IEEE*, vol. 86, no. 1, pp. 82–85, Jan. 1998.
- [5] M. Nazirzadeh, M. Shamsabardeh, and S. J. Ben Yoo, “Energy-efficient and high-throughput nanophotonic neuromorphic computing,” in *Proc. Conf. Lasers Electro-Opt.*, 2018, Paper ATH3Q.2.
- [6] R. H. Dennard *et al.*, “Design of ion-implanted mosfets with very small physical dimensions,” *IEEE J. Solid-State Circuits*, vol. SC 9, no. 5, pp. 256–268, Oct. 1974.
- [7] R. Kurzweil, *The Singularity is Near: When Humans Transcend Biology*. Baltimore, MD, USA: Penguin, 2006.
- [8] S. Borkar and A. A. Chien, “The future of microprocessors,” *Commun. ACM*, vol. 54, no. 5, pp. 67–77, 2011.
- [9] A. W. Topol *et al.*, “Three-dimensional integrated circuits,” *IBM J. Res. Develop.*, vol. 50, no. 4.5, pp. 491–506, 2006.
- [10] T. Coughlin and J. Handy, “Emerging memory poised to explode,” 2018. [Online]. Available: <https://www.snia.org/sites/default/files/SSSI/Emerging%20Memory%20Poised%20to%20Explode%20final%20webcast%2012-11-18.pdf>
- [11] D. M. Gill *et al.*, “Demonstration of error free operation up to 32 Gb/s from a CMOS integrated monolithic nano-photonic transmitter,” in *Proc. Conf. Lasers Electro-Opt.*, 2015, pp. 1–2.
- [12] C. Sun *et al.*, “Single-chip microprocessor that communicates directly using light,” *Nature*, vol. 528, pp. 534–538, 2015.
- [13] K. Shang, S. Pathak, B. Guan, G. Liu, and S. J. B. Yoo, “Low-loss compact multilayer silicon nitride platform for 3D photonic integrated circuits,” *Opt. Express*, vol. 23, no. 16, pp. 21334–21342, 2015.
- [14] K. Shang *et al.*, “Tri-layer, vertical Y-junction, Si₃N₄/SiO₂ 3D photonic integrated circuits with arbitrary splitting ratio,” presented at the Opt. Fiber Commun. Conf., 2016.
- [15] W. D. Sacher *et al.*, “Tri-layer silicon nitride-on-silicon photonic platform for ultra-low-loss crossings and interlayer transitions,” *Opt. Express*, vol. 25, no. 25, pp. 30862–30875, 2017.
- [16] AIM_Photronics. 2019. [Online]. Available: <http://www.aimphotonics.com/>
- [17] W. A. Vitale *et al.*, “Ultra fine-pitch TSV technology for ultra-dense high-Q RF inductors,” in *Proc. Symp. VLSI Technol.*, pp. T52–T53.
- [18] A. Dembla, Y. Zhang, and M. S. Bakir, “Fine pitch TSV integration in silicon micropin-fin heat sinks for 3D ICs,” in *Proc. IEEE Int. Interconnect Technol. Conf.*, pp. 1–3.
- [19] Y. Zhang, Y. Ling, Y. Zhang, K. Shang, and S. J. B. Yoo, “High-density wafer-scale 3-D silicon-photonic integrated circuits,” *IEEE J. Sel. Topics Quantum Electron.*, vol. 24, no. 6, Nov./Dec. 2018, Art. no. 8200510.
- [20] H. W. Chen, Y. H. Kuo, and J. E. Bowers, “A hybrid silicon-AlGaInAs phase modulator,” (in English), *IEEE Photon. Technol. Lett.*, vol. 20, no. 21–24, pp. 1920–1922, Nov./Dec. 2008.
- [21] C. R. Doerr and K. Okamoto, “Advances in silica planar lightwave circuits,” *J. Lightw. Technol.*, vol. 24, no. 12, pp. 4763–4789, Dec. 2006.
- [22] R. Won, “Silicon photonics: Low-loss and compact,” *Nature Photon. Res. Highlights*, vol. 9, no. 10, pp. 631–631, 2015.
- [23] M. J. R. Heck, J. F. Bauters, M. L. Davenport, D. T. Spencer, and J. E. Bowers, “Ultra-low loss waveguide platform and its integration with silicon photonics,” *Laser Photon. Reviews*, vol. 8, no. 5, pp. 667–686, 2014.
- [24] Y. Zhang *et al.*, “Foundry-enabled scalable all-to-all optical interconnects using silicon nitride arrayed waveguide router interposers and silicon photonic transceivers,” *IEEE J. Sel. Topics Quantum Electron.*, vol. 25, no. 5, Sep./Oct. 2019, Art. no. 8300409.
- [25] Y. Zhang *et al.*, “Sub-wavelength spacing optical phase array nanoantenna emitter with vertical silicon photonic vias,” in *Proc. Opt. Fiber Commun. Conf. Expo.*, 2018, Paper Tu2A.6.
- [26] K. P. Shang *et al.*, “Uniform emission, constant wavevector silicon grating surface emitter for beam steering with ultra-sharp instantaneous field-of-view,” (in English), *Opt. Express*, vol. 25, no. 17, pp. 19655–19661, Aug. 21, 2017.
- [27] Y. Zhu *et al.*, “Fabrication of keyhole-free ultra-deep high-aspect-ratio isolation trench and its applications,” *J. Micromechanics Microengineering*, vol. 15, no. 3, pp. 636–642, 2005.
- [28] D. A. Miller, “Attojoule optoelectronics for low-energy information processing and communications—A tutorial review,” *J. Lightw. Technol.*, vol. 35, no. 3, pp. 346–396, 2017.
- [29] X. Sheng *et al.*, “Transfer printing of fully formed thin-film microscale GaAs lasers on silicon with a thermally conductive interface material,” *Laser Photon. Rev.*, vol. 9, no. 4, pp. L17–L22, 2015.
- [30] E. M. Izhikevich, “Simple model of spiking neurons,” *IEEE Trans. Neural Netw.*, vol. 14, no. 6, pp. 1569–1572, Nov. 2003.
- [31] P. Diehl and M. Cook, “Unsupervised learning of digit recognition using spike-timing-dependent plasticity,” (in English), *Frontiers Comput. Neuroscience, Original Res.*, vol. 9, no. 99, pp. 1–9, 2015.
- [32] T. Su *et al.*, “Interferometric imaging using Si₃N₄ photonic integrated circuits for a SPIDER imager,” *Opt. Express*, vol. 26, no. 10, pp. 12801–12812, 2018.
- [33] T. Su *et al.*, “Experimental demonstration of interferometric imaging using photonic integrated circuits,” *Opt. Express*, vol. 25, no. 11, pp. 12653–12665, 2017.
- [34] D. Liang *et al.*, “150 mm InP-to-silicon direct wafer bonding for silicon photonic integrated circuits,” *ECS Trans.*, vol. 16, no. 8, pp. 235–241, 2008.
- [35] S. Chen *et al.*, “Electrically pumped continuous-wave III–V quantum dot lasers on silicon,” *Nature Photon.*, vol. 10, pp. 307–311, 2016.
- [36] L. Sanchez *et al.*, “Collective die direct bonding for photonic on silicon,” *ECS Trans.*, vol. 86, no. 5, pp. 223–231, Jul. 20, 2018.
- [37] B. Szelag *et al.*, “Hybrid III–V/silicon technology for laser integration on a 200-mm fully CMOS-compatible silicon photonics platform,” *IEEE J. Sel. Topics Quantum Electron.*, vol. 25, no. 5, Sep./Oct. 2019, Art. no. 8201210.
- [38] A. Carlson, A. M. Bowen, Y. Huang, R. G. Nuzzo, and J. A. Rogers, “Transfer printing techniques for materials assembly and micro/nanodevice fabrication,” *Adv. Mater.*, vol. 24, no. 39, pp. 5284–5318, 2012.
- [39] M. A. Meitl *et al.*, “Transfer printing by kinetic control of adhesion to an elastomeric stamp,” *Nature Mater.*, vol. 5, no. 1, pp. 33–38, 2006.
- [40] C. A. Bower, M. A. Meitl, S. Bonafede, and D. Gomez, “Heterogeneous integration of microscale compound semiconductor devices by micro-transfer-printing,” in *Proc. IEEE 65th Electron. Compon. Technol. Conf.*, 2015, pp. 963–967.
- [41] J. Zhang *et al.*, “Transfer-printing-based integration of a III–V-on-silicon distributed feedback laser,” *Opt. Express*, vol. 26, no. 7, pp. 8821–8830, 2018.
- [42] J. Justice *et al.*, “Wafer-scale integration of group III–V lasers on silicon using transfer printing of epitaxial layers,” *Nature Photon.*, vol. 6, pp. 610–614, 2012.
- [43] N. Ye *et al.*, “Transfer print integration of waveguide-coupled germanium photodiodes onto passive silicon photonic ICs,” *J. Lightw. Technol.*, vol. 36, no. 5, pp. 1249–1254, Mar. 2018.
- [44] J. McPhillimy *et al.*, “High accuracy transfer printing of single-mode membrane silicon photonic devices,” *Opt. Express*, vol. 26, no. 13, pp. 16679–16688, 2018.

Yu Zhang received the B.S. degree in optoelectronics from the Huazhong University of Science and Technology, Wuhan, China, in 2010 and the Ph.D. degree in electronic and computer engineering from The Hong Kong University of Science and Technology, Hong Kong, in 2016. He is currently an Assistant Project Scientist in electrical and computer engineering with The University of California, Davis, Davis, CA, USA. His current research interests focus on 3-D integrated silicon photonic optical phased array, hybrid silicon laser, amplifier and phase modulators, and silicon photonic integrated circuit enabled all-to-all interconnection for scalable high-performance computing systems.

Anirban Samanta received the B.E. degree in electronics and communications engineering from the Birla Institute of Technology, Mesra, Mesra, India, in 2012 and the M.S. degree from The University of California Irvine, Irvine, CA, USA, in 2014. He is currently working toward the Ph.D. degree in electrical and computer engineering with The University of California, Davis, Davis, CA. He has previously worked with the ARIANNA Neutrino Detector Group at Irvine, developing sensor systems for neutrino detection.

His current research interests focus on electronic-photonic co-simulation, co-design and integration with applications in high performance computing and data centers.

S. J. Ben Yoo (Fellow, IEEE) received the B.S. degree in electrical engineering with distinction, the M.S. degree in electrical engineering, and the Ph.D. degree in electrical engineering with a minor in physics, from Stanford University, Stanford, CA, USA, in 1984, 1986, and 1991, respectively. He is currently a Distinguished Professor of electrical engineering with UC Davis, Davis, CA, USA. His research with UC Davis includes 2D/3D photonic integration for future computing, communication, imaging, and navigation systems, micro/nano systems integration, and the future Internet. Prior to joining UC Davis in 1999, he was a Senior Research Scientist with Bellcore, leading technical efforts in integrated photonics, optical networking, and systems integration. His research activities with Bellcore included the next-generation internet, reconfigurable multiwavelength optical networks (MONET), wavelength interchanging cross connects, wavelength converters, vertical-cavity lasers, and high-speed modulators. He led the MONET testbed experimentation efforts, and participated in ATD/MONET systems integration and a number of standardization activities. Prior to joining Bellcore in 1991, he conducted research on nonlinear optical processes in quantum wells, a four-wave-mixing study of relaxation mechanisms in dye molecules, and ultrafast diffusion-driven photodetectors with Stanford University. He is a fellow of OSA, NIAC, and was the recipient of the DARPA Award for Sustained Excellence (1997), the Bellcore CEO Award (1998), the Mid-Career Research Faculty Award (2004 UC Davis), and the Senior Research Faculty Award (2011 UC Davis).

Kuanping Shang received the B.S. degree in physics with a second bachelor's degree in economics from Peking University, Beijing, China, in 2011 and the Ph.D. degree in electronic and computer engineering from The University of California, Davis, Davis, CA, USA, in 2017. He is currently working with Intel. From June 2009 to 2010, he was a Research Assistant with the Institute of Modern Optics, Peking University. From June 2010 to 2011, he was a Research Assistant with the State Key Lab for Mesoscopic Physics, China.