Fast and efficient Sb-based type-II phototransistors integrated on silicon

Cite as: APL Photon. 10, 036106 (2025); doi: 10.1063/5.0233887
Submitted: 18 August 2024 • Accepted: 12 February 2025 •
Published Online: 3 March 2025

Lining Liu,[1] Simone Bianconi,[1] Skyler Wheaton,[1] Nathaniel Coirier,[1] Farah Fahim,[2] and Hooman Mohseni[1,a)]

AFFILIATIONS

[1] Bio-Inspired Sensors and Optoelectronics Laboratory, Northwestern University, 2145 Sheridan Rd, Evanston, Illinois 60208, USA
[2] ASIC Development Group, Particle Physics Division, Fermi National Accelerator, Batavia, Illinois 60510, USA

[a)] **Author to whom correspondence should be addressed:** hmohseni@northwestern.edu

ABSTRACT

Increasing the energy efficiency and reducing the footprint of on-chip photodetectors enable dense optical interconnects for emerging computational and sensing applications. While heterojunction phototransistors (HPTs) exhibit high energy efficiency and negligible excess noise factor, their gain-bandwidth product (GBP) has been inferior to that of avalanche photodiodes at low optical powers. Here, we demonstrate that utilizing type-II energy band alignment in an Sb-based HPT results in six times smaller junction capacitance per unit area and a significantly higher GBP at low optical powers. These type-II HPTs were scaled down to 2 μm in diameter and fully integrated with photonic waveguides on silicon. Thanks to their extremely low dark current and high internal gain, these devices exhibit a GBP similar to the best avalanche devices (~270 GHz) but with one order of magnitude better energy efficiency. Their energy consumption is about 5 fJ/bit at 3.2 Gbps, with an error rate below 10^{-9} at −25 dBm optical power at 1550 nm. These features suggest new opportunities for creating highly efficient and compact optical receivers based on phototransistors with type-II band alignment.

I. INTRODUCTION

Integrated photodetectors have attracted great interest for a wide variety of applications, including photonic neural network accelerators,[1,2] photonic signal processors,[3] 3D depth imaging,[4] and optical interconnects.[5–7] In particular, optical interconnects are considered to be one of the most promising candidates for next-generation on-chip communication[7] thanks to their unique advantages.[8–10] However, current on-chip implementations fail to meet the footprint and energy consumption requirements necessary to surpass the performance of electrical interconnects.[7,11] Energy efficiency and energy-per-bit are key figures of merit that limit the growth in data storage, transmission, and computation; hence, improving both these metrics in key components is essential.[11–14] In this work, we focus on the receiving end of the on-chip data transmission and present a new high-performance integrated optical receiver with significant improvements in both energy efficiency and energy-per-bit over the existing optical receivers integrated on silicon. In order to drive meaningful on-chip loads with the minimum input optical signal, an integrated receiver requires some form of amplification of the detected signal. Detectors without internal gain mechanisms (e.g., p–i–n photodiodes) rely entirely on external amplifiers; however, even the latest generations of custom-designed high-speed complementary metal–oxide–semiconductor (CMOS) amplifiers[15] require significant power and on-chip area (see the supplementary material Table S3). Conversely, most state-of-the-art optical receivers, such as avalanche photodiodes (APDs) and heterojunction phototransistors (HPTs), utilize internal amplification to boost the detected signal in order to either directly drive a load or to reduce the complexity of the external amplifiers.[16,17] In this configuration, higher internal gains enable to either directly drive larger loads or to eliminate the need for external amplifiers. The large excess noise factor of APDs limits their capability in operating at high gain; hence, external amplifiers are typically still necessary; conversely, HPTs can achieve large internal amplification at high speed with relatively low excess noise and dark current. Notably, staircase APDs have achieved a low excess noise factor,[18] but the maximum achievable gain has remained small since the

invention of these devices about four decades ago.[19] Early studies suggested that HPTs could be used for high-speed receivers,[20] and by the year 2000, they achieved bitrates of 40 Gbps.[21] Unfortunately, these high-speed HPT receivers suffered from low sensitivity (needed about −8 dBm optical power), and they were not integrated on silicon. These issues have moved the researchers' focus toward APD receivers in the past decade.[9,22–24] However, besides the limitations on gain discussed above, most CMOS-compatible APDs also require a large bias voltage to achieve proper gain values and exhibit high dark current,[9,17,25–27] leading to poor energy efficiencies and large power consumption. Here, we demonstrate that using a type-II energy band alignment in a properly designed HPT leads to devices that can simultaneously achieve a high gain-bandwidth product (GBP) at low optical powers, high sensitivity, and high energy efficiency. Our type-II heterojunction phototransistor (T2-HPT) integrated on silicon achieves a GBP similar to that of the best reported APDs on silicon, but with much lower dark current and operating bias voltage than APDs, a much smaller footprint that enables dense integration, and about ten times higher energy efficiency for a given load and similar maximum bitrate. More importantly, we show that it achieves such a high speed at a low optical power—well below that of the best HPTs previously reported.[28–31] Our experimental and modeling results suggest that the superior performance of these devices is due to the lower capacitance of type-II heterojunctions compared to the traditional type-I heterojunctions, as well as the better thermal management of our integration method. All results presented here are based on devices integrated on silicon and optically coupled to on-chip hydrogenated amorphous silicon (a-Si:H) waveguides and grating couplers using a CMOS-compatible process (see Sec. V). Since the integration approach is not limited to a specific CMOS process and the optical interconnects (i.e., waveguides and couplers) are not substrate-specific, our approach can be used with most existing CMOS technologies.

II. RESULTS

A. Device design

The detailed epitaxial structure of the HPT detector wafers, grown on InP substrates by low-pressure metal–organic chemical vapor deposition (LP-MOCVD), is shown in Table S1 in the supplementary material. The main bi-polar junction phototransistor layers consist of an n-doped InGaAs collector layer and an n-doped InP emitter layer, separated by a 50-nm thick p-doped GaAsSb base layer. The GaAsSb base layer employed in this work is specifically designed to leverage the type-II band alignment between GaAsSb and InGaAs/InP to achieve a high GBP, as discussed in more detail in Secs. II B–V.

In order to fabricate the integrated optical receivers, we used a low-temperature Mo/Au–Au/Mo bonding method[32] to transfer epitaxial layers to silicon substrates, as detailed in the supplementary material and depicted in Fig. 1(a). We observed a typical root-mean-square (rms) roughness of ∼2 nm over large areas of the transferred film, indicating high-quality bonding (see supplementary material Sec. 1). After the transfer, T2-HPTs with different sizes (ranging from 100 to 2 μm in diameter) were fabricated on the silicon substrate and integrated with a-Si:H waveguides and grating couplers (Fig. 1). a-Si:H can be easily deposited at low temperatures (≤300 °C) in a process compatible with III–V materials and back end of line (BEOL) CMOS processes.[33,34] Our approach allows fine control of the a-Si:H refractive index by tuning the RF power and pressure during deposition[35] (see Sec. V and supplementary material Sec. 1). Figure 1(b) shows a scanning electron microscope (SEM) image of a T2-HPT device before the waveguide integration, with clear collector, base, and emitter layers of the HPT. Figure 1(c) offers a cross-sectional SEM image of a detector after waveguide integration showing the interface between the waveguide and the detectors.

B. Device measurement

We characterized the fully integrated T2-HPT devices with the architecture shown in Fig. 1(d). While we measured devices with different diameters, here, we mainly focus on the performance of devices with a 2 μm diameter. This is because the speed of these T2-HPT photodetectors tends to increase with decreasing device diameter due to the corresponding decrease in junction capacitance, as shown in part 3 of the supplementary material and discussed in detail in Secs. III–V. In addition, the smaller size of these devices comes with the added benefit of a reduced footprint for optical interconnects. We measured the dark current (I_d) and the photo-current (I_{ph}) of these devices at room temperature by coupling the free-space collimated output of a 1.55 μm laser diode into the grating couplers [see Fig. 1(d) and Fig. S1(a) of the supplementary material]. Figure 2(a) shows the dark and photo-current of a 2 μm T2-HPT. The device dark current at 2 V is about 0.5 nA, which is several orders of magnitude lower than the dark current of the best integrated APDs at their operating bias voltages. This is despite the dark current density increasing when shrinking the device diameters under 10 μm, due to a significant contribution from surface effects, which suggests that dark currents could be reduced with better surface passivation (see supplementary material Sec. 2). The photo-current shown in Fig. 2(a) was measured at an optical power of −19 dBm (12.6 μW) coupled to the 2 μm T2-HPT. Figure 2(b) shows the DC responsivity of T2-HPT devices of different diameters at −19 dBm coupled optical power: the responsivity of all devices exceed 100 A/W at voltages as low as 0.75 V. Notably, devices smaller than 5 μm have lower responsivity, most likely due to the reduced overlap with the optical mode of the 5 μm-wide waveguides. This effect could be prevented by optimizing the width of the waveguides or introducing tapered waveguides. Furthermore, the low coupling to small devices can be addressed by employing hybrid optical antennas in order to enhance the local density of states and coupling of the near-field to the waveguide modes.[36–38]

We evaluated the performance of the 2 μm T2-HPT as a digital receiver, using bit error rate (BER) measurements (see Sec. V). We coupled the output beam of a 1.55 μm laser diode, which was directly modulated by pseudo-random binary sequences (PRBS), to the integrated grating couplers. Without any external amplifier, we achieved high-quality open eye diagrams at 3.2 Gbps—the limit of our bit error rate test (BERT) system—as shown in Fig. 2(c). We measured sensitivities of −26 and −25 dBm at 2.5 and 3.2 Gbps, respectively, as defined for a $BER < 10^{-9}$ [Fig. 2(d)].[23] More importantly, the high gain and direct driving capability of our device eliminated the need for a high-gain and low-noise amplifier in these experiments, and we directly coupled our device to the 50-Ω

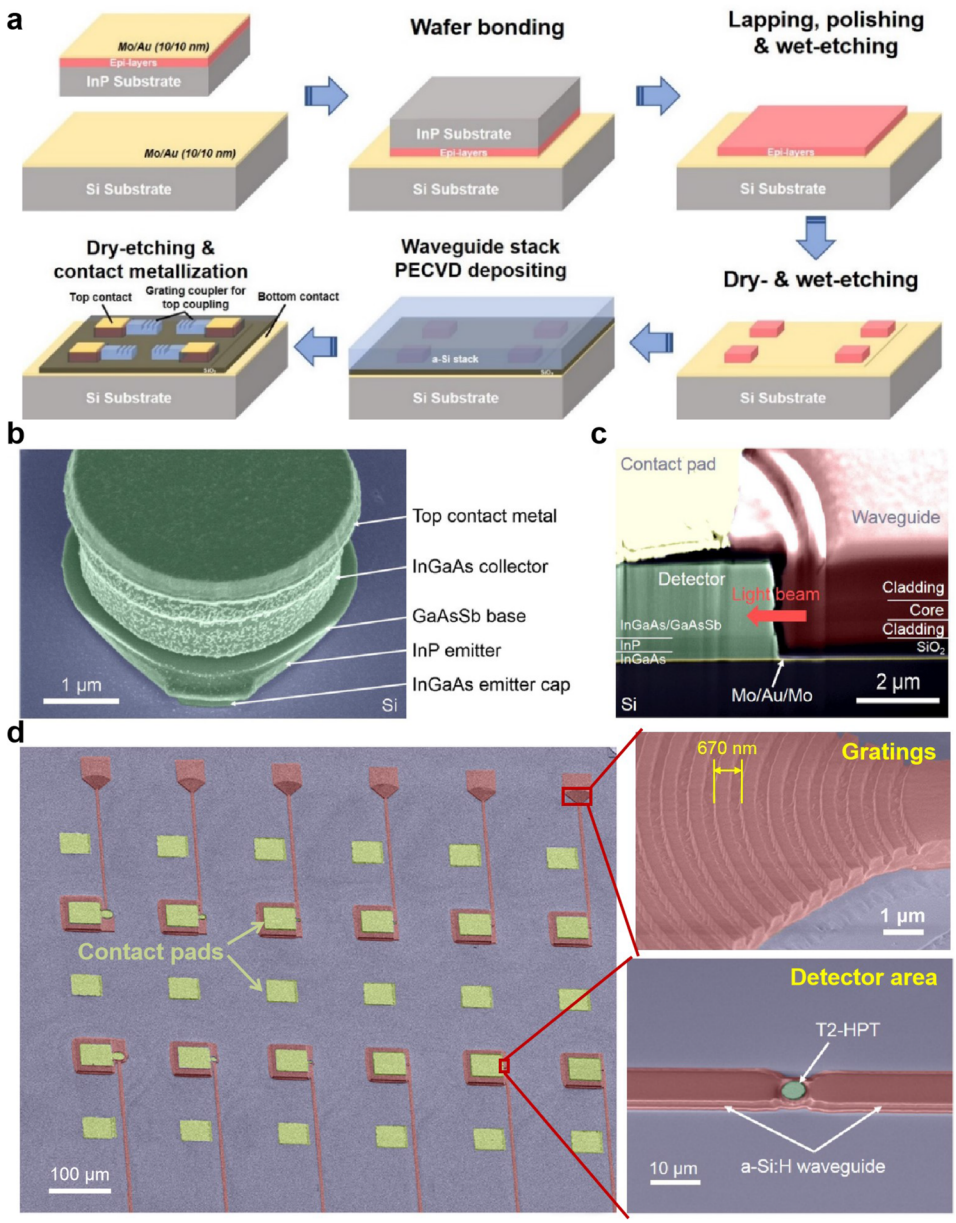

FIG. 1. Integrated T2-HPT presented in this work. (a) Fabrication process of the integrated T2-HPT detector. (b) The SEM image of the detector after pillar etching shows the epitaxial III–V semiconductor layers on the silicon substrate. (c) Cross-sectional FIB-SEM image of an integrated device at the end of the fabrication process. (d) SEM image of the integrated detectors with different sizes, together with zoomed-in views of the grating coupler and the detector areas.

load. Since our BERT system was limiting our measured data rate to 3.2 Gbps, we evaluated the actual bandwidth and the resulting estimated bitrate, using a fast optical pulse and a fast oscilloscope (see Sec. V). Figure 2(e) shows the full width at half maximum (FWHM) of the output signal of a 2 μm T2-HPT when receiving an optical pulse with a FWHM of ~7 ps. The electrical signal on a 50 Ω load shows a FWHM of 97 ps and a jitter of less than 9 ps. The corresponding 3 dB bandwidth is about 5 GHz, based on the Fourier transform of the output pulse.[16] This value is in close agreement with our simulation (see Sec. 3 of the supplementary material). The measured SNR at the output pulse is about 22.5, suggesting a data rate of 10 Gbps at $BER < 10^{-9}$ (well beyond the bitrate limit of our BERT system). Even using the measurement-limited data rate of 3.2 Gbps, the 2 μm detector with no external amplifier represents a data rate density of 800 Tbps/mm^2, which is significantly higher than that of the best reported amplifier-free APDs (~100 Tbps/mm^2 in Ref. 9). A high data rate density is crucial for the future on-chip optical interconnects,[11] enabling the next-generation computing and sensing chips with massive data bandwidths.

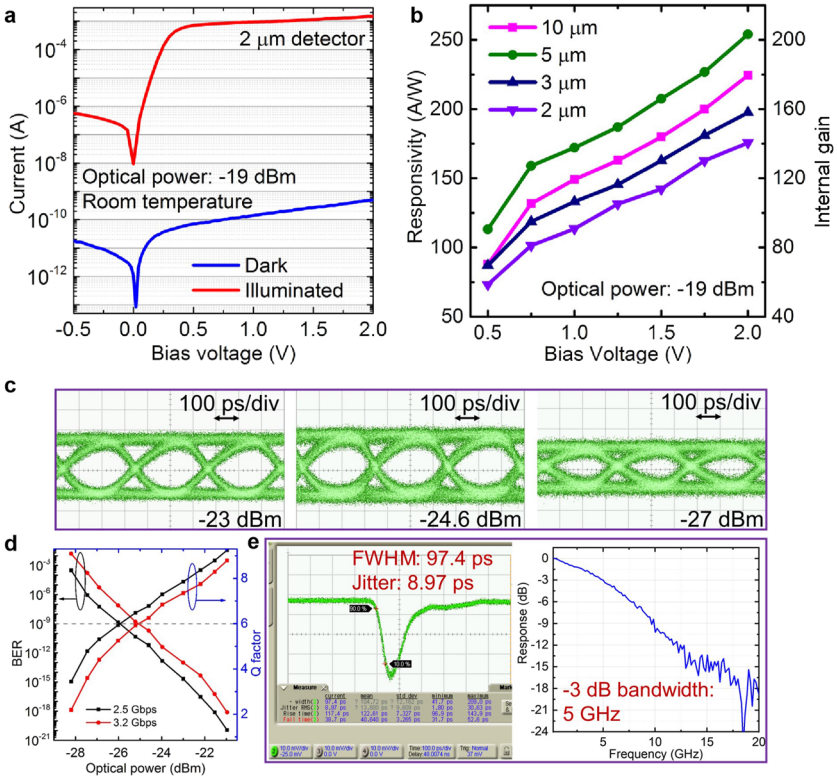

FIG. 2. Characterization measurements of the integrated T2-HPT with optical signal coupled through the grating coupler. (a) Current–voltage characteristic of the integrated 2 μm detectors at room temperature in the dark and under illumination. (b) DC responsivity R and internal gain of T2-HPT detectors of different sizes as a function of bias voltage at an optical power of −19 dBm (12.6 μW) coupled to the detector. (c) Eye diagrams of 2 μm device at a 3.2 Gbps data rate, with different optical powers coupled to the detector. (d) Bit error rate (BER) and Q factor as a function of optical power, extracted from the eye diagrams of the 2 μm detector. (e) Output pulse from the integrated 2 μm detector and its Fourier transform.

III. DISCUSSION

It is known that type-II band alignment helps preventing the onset of the base push-out at high current densities, known as the Kirk effect.[39] However, we hypothesized that type-II band alignment can also be used to decrease the device junction capacitance significantly. Reducing junction capacitance of HPTs not only makes the phototransistors faster but also makes them more sensitive, as we had proposed[40] and recently demonstrated experimentally.[41] Figure 3 presents the difference between type- I and type-II band alignment: the discontinuity of the type-II alignment provides a clear advantage from the charge storage point of view, since it prevents excess holes from traveling into the collector and precluding the onset of the Kirk effect.[29] In addition, a type-II band alignment helps reducing the formation of an electrostatic barrier inside the collector, resulting in a slower increase in capacitance at increased current levels. To evaluate this hypothesis, we measured the capacitance of the type-II HPT and a type-I HPT with equal base doping levels, thicknesses, and diameters. Our experimental results show that type-II HPT exhibits about six times smaller capacitance per unit area than type-I HPT at a bias voltage of 2 V utilized in this work (0.3 fF/μm^2 for the type-II structure vs 1.8 fF/μm^2 for the type-I structure; see Fig. 3(c). The measured capacitance values and their bias dependencies are in good agreement with numerical simulations for type-II devices [see Fig. 3(d)].

In addition to the reduced capacitance, our devices are made with special attention to the high current density needed to directly drive loads without an amplifier. Numerical simulations show that compared to our previously reported devices on a native substrate, the bonding method and the small dimensions of the devices reported in this work significantly enhance their thermal conductance and reduce their internal temperature (see supplementary material Sec. 4 for thermal modeling). At an output power of about 2 mW (2 V bias voltage and 1 mA photo-current), the temperature increase in the transferred detector is more than four times lower than that in the as-grown detector. A high temperature can degrade the GBP of the device, mainly due to the increased base transit time,[42] as previously demonstrated for type-I and type-II HBTs.[43,44] A direct consequence of a fundamentally lower junction capacitance and operating voltage is better energy efficiency. While the presented device is not truly optimized, we would like to compare its energy efficiency with some of the best reported APD and HPT devices. The energy efficiency of an optical receiver can be calculated as $\eta = E_U/(E_C + E_U)$, where E_C is the energy waste per bit and E_U is the energy delivered to the load per bit [see Fig. 4(a)]. For approaches based on optical receivers in CMOSs, these values can be estimated as (see supplementary material Sec. 5)

$$E_U \approx \frac{1}{2} C_L V_L^2 \tag{1}$$

$$E_C \approx \frac{2V_b I_d + V_b I_{ph} + V_L I_d + V_L I_{ph}/2}{2BR}, \tag{2}$$

where C_L and V_L are the load capacitance and voltage (around 1 V in modern CMOS), respectively; I_d, I_{ph}, and V_b are the dark current,

photo-current, and voltage bias of the photodetector, respectively; and BR is the bitrate. From these equations, it is evident that V_b and I_d are the most important factors that determine the energy consumption and efficiency of the receiver. As mentioned above, existing CMOS-compatible APDs require a large V_b to achieve a reasonable avalanche gain and exhibit a large I_d due to the large electric field required for avalanche, both of which lead to poor energy efficiencies mostly around 1%.

In Fig. 4, we compare the GBP vs energy efficiency of the best reported results from optical receivers, including III–V HPTs and CMOS-compatible APDs.[9,17,23,24,26,28–31,45–48] It is worth noting that the best HPT receivers included in this comparison are now over a decade old: most recent integrated phototransistor detectors, including both FET and HPT, are focused on rather high sensitivity but low speed[49–54] and hence do not exhibit relevant performance for the comparison drawn here. Figure 4(b) shows the GBP vs energy efficiency η and consumed energy per bit E_C (details in supplementary material Sec. 6). Compared to integrated APDs, our integrated T2-HPT exhibits about one order of magnitude higher energy efficiency $\eta \sim 17\%$ and a significantly lower energy consumption, $E_C \sim 5$ fJ/bit. Most reports have shown integrated APDs exceeding 10 Gbps bitrates using amplifiers. However, the addition of amplifier circuits has limited the ability to achieve the longstanding goals of high energy efficiency and high data rate densities. Without any amplifier, the maximum achievable bitrate for a given load is set by the responsivity and photo-current of the device and hence its sensitivity. Since the excess noise factor of APDs grows with their internal gain, the practical responsivity of APDs has been limited to below ~30 A/W. Therefore, without an amplifier, the achievable bitrate of APDs would be comparable to that of our device (see detailed discussion in supplementary material Sec. 6 and Table S4). For applications that require a large number of channels per area but moderate bitrates per channel, our device could be used without an amplifier to achieve a very energy efficient solution within an extremely small physical footprint.

Furthermore, for applications that require high bitrates per channel, the high GBP of the T2-HPT devices still allows use of simpler, more compact, and more efficient amplifiers. As an example, we evaluated the performance of these devices when integrated with a compact 5-transistor CMOS amplifier (see supplementary material Sec. 7 for details on the amplifier). Using the compact amplifier, the receiver could achieve a bitrate of 16.7 Gbps [Fig. S8(b)]. When including capacitive loads of 5 and 50 fF, the highest open-eye data rates were ~15 Gbps [Fig. S8(c)] and ~10 Gbps [Fig. S8(d)], respectively. Crucially, the compact amplifier (supplementary material Sec. 7) adds an energy dissipation of ~60 fJ/bit, which is smaller than that of typical amplifiers used in other integrated receivers with a similar CMOS technology node, while achieving approximately ten times higher sensitivity and about two times larger bitrate.[55]

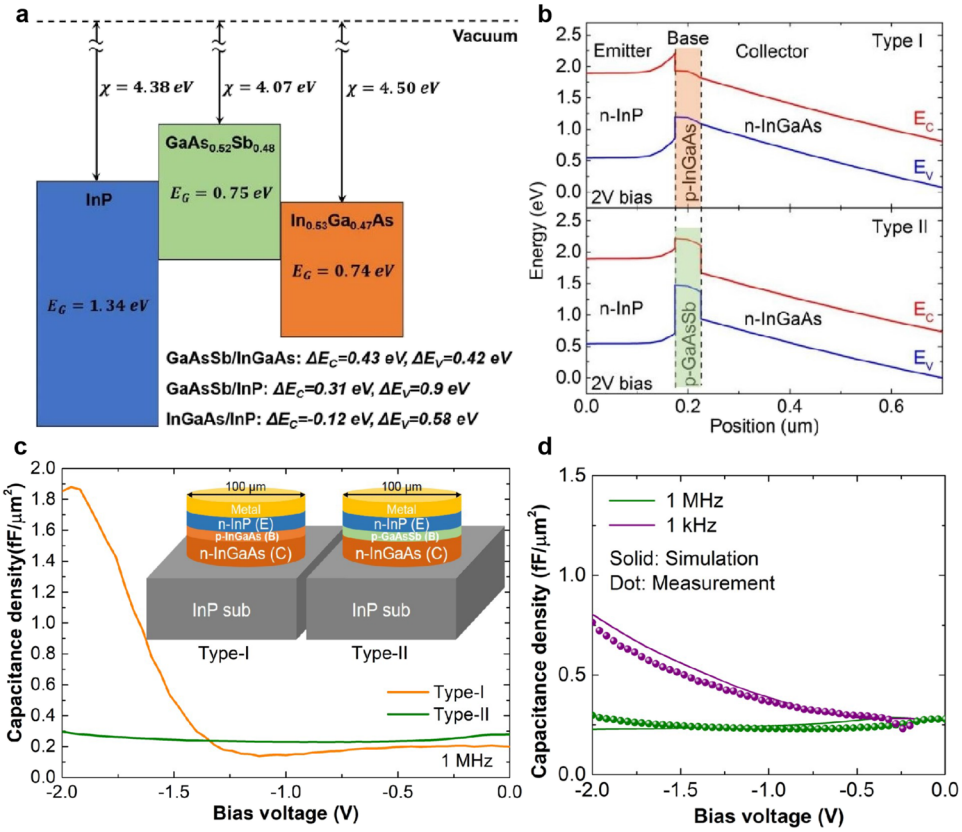

FIG. 3. Band structure of type-I and type-II. (a) Schematic diagram of the energy band alignment of the InGaAs, GaAsSb, and InP alloy compositions used in this work. GaAsSb forms a type-II alignment with InGaAs and InP, while InGaAs/InP forms a type-I alignment. (b) Simulated band structures of two NPN HPTs based on type-I (top) and type-II (bottom) band alignments. (c) Experimentally measured junction capacitance vs voltage (C–V) for type-I and type-II HPTs with identical layer thicknesses and doping levels show that type-II HPT has a substantially lower junction capacitance at an operating bias of −2 V (d) Simulated C–V for the type-II HPTs structure, showing good agreement with experimental measurements at both low and high frequencies.

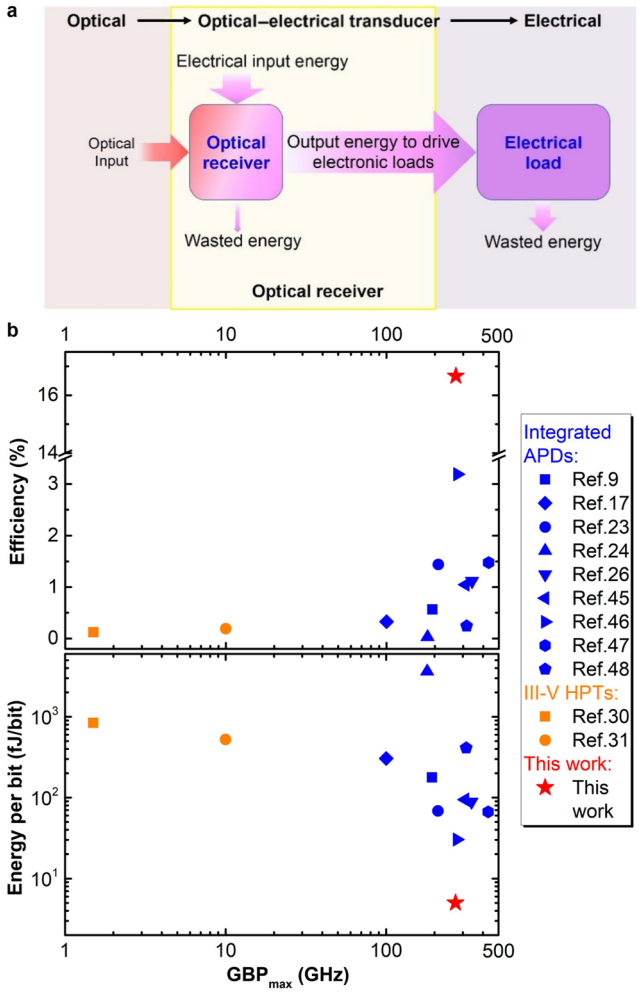

FIG. 4. Energy consumption of integrated optical receivers. (a) The optical receiver transduces the input optical signal to an electrical signal for on-chip data transmission. The input energy of the system consists of an optical signal and an electrical bias, part of which is wasted by the dark current of the optical receiver and the rest of which is transduced to the electrical load. At the electrical load, part of the energy is utilized, while the other part is wasted. The calculation of the values of these energies is discussed in the text. (b) Comparison of this work with the best reported detectors for optical receivers. Top: GBPs vs energy efficiency (η). Note that we removed a portion of the y axis, from 2% to 12%, to increase the visibility. Bottom: GBPs vs total consumed energy (E_C) of detectors at an internal gain of 15 and 2 fF load for an optical power of about −30 dBm (see detailed calculation and tabulated data in supplementary material Sec. 6).

IV. CONCLUSION

We demonstrated a type-II heterojunction phototransistor (T2-HPT) with exceptional performance characteristics. Arrays of devices with different sizes were integrated on silicon wafers using a CMOS-compatible wafer bonding and an additive waveguide interconnect method. Electrical and optical characterizations show that the integrated T2-HPT can achieve a gain-bandwidth product of ~270 GHz at ~10 μW optical power. This device is the first HPT that simultaneously shows a high speed and high sensitivity. When used as an optical receiver, with and without an amplifier, the device showed bitrate and sensitivity values that are similar to those of the best APDs. However, the T2-HPT showed about ten times higher energy efficiency and significantly lower energy consumption per bit compared with APDs, due to its extremely low dark current and operating voltage. Our experimental and simulation results support the hypothesis that the type-II band alignment is potentially the reason for achieving exceptionally high responsivity and low junction capacitance in our device. In parallel, these tiny devices show very large internal gain and current densities required for directly driving large capacitive loads, leading to a massive data transmission density in excess of 800 Tbps/mm^2 at an attractive energy consumption of ~5 fJ/bit. The unique combination of a small footprint, low energy consumption per bit, and high energy efficiency makes these new devices a promising choice for high-density optical receivers. We hope that our results encourage the research community to further evaluate type-II phototransistors based on new material systems such as van der Waals heterostructures,[56] which present an excellent opportunity for application of this strategy to a broad set of materials and wavelengths.

V. MATERIALS AND METHODS

A. Fabrication process

The T2-HPT detectors used in this work are grown by metalorganic chemical vapor deposition (MOCVD) on n-doped InP substrates. Figure 1(a) and supplementary material Sec. 2 describe the device fabrication process flow, which starts with the degreasing of a ~1 × 1 cm^2 sample and a ~1.7 × 1.7 cm^2 silicon substrate in organic solvents. Then, NH_4OH and $(NH_4)_2S$ treatment is performed on the T2-HPT sample to remove the native oxide and passivate the surface; meanwhile, the silicon substrate is dipped in buffered oxide etchant (BOE) for native oxide removal. Both samples are then immediately transferred to the chamber of an electron-beam evaporator, followed by the evaporation of the Mo/Au (10/10 nm) bilayer. After an Ar-based plasma treatment is used for surface activation, the two samples are bonded to each other by a wafer-bonding tool (FC-150 by SET) at 200 °C and with a force of 35 kg. Then, HCl (37%) is used to remove the InP substrate and leave behind the epitaxially grown layers. Subsequently, the T2-HPT devices are defined by conventional photolithography and a two-step etching process (dry followed by wet etching). To passivate the sidewalls of the detectors, SiO_2 (300 nm) is deposited by Plasma Enhanced Chemical Vapor Deposition (PECVD) at 300 °C. Keeping the sample inside the PECVD chamber, the hydrogenated amorphous silicon (a-Si:H) waveguide stack is then deposited. It consists of a 400 nm lower cladding layer (RF 40 W, 900 mTorr), a 300 nm core layer (RF 50 W, 300 mTorr), and an 800 nm upper cladding layer (RF 40 W, 900 mTorr). Waveguides with a width of 5 μm and grating couplers with a period of 670 nm are etched to the lower cladding layer using reactive-ion etching (RIE) and inductively coupled plasma etching (ICP), respectively. Benzocyclobutene (BCB) is then spin-coated, cured, and etched back for passivation and planarization. Finally, the Ti/Ni/Au (20/30/100 nm) top and bottom contacts are evaporated after the etching of openings. The temperature of the whole fabrication process never exceeds 300 °C.

B. Characterization and measurement

The electrical and optical DC performance of the fabricated HPTs are characterized by utilizing a 1.55-μm laser diode and a digital multimeter (34410A), connected with a low-noise current preamplifier (SR 570) at room temperature. The light is delivered to the detectors through the waveguide by focusing the beam to the grating coupler. For laser power calibration, the incident optical power through the objective lens was measured using a commercial InGaAs-based p–i–n photodiode module in a dark environment. Pulse measurement is conducted using a Calmar Optcom Femtosecond pulse laser as the light source emitting an optical pulse with a full width at half maximum (FWHM) of 7.31 ps. The output of the device is probed with a GSG probe and directly measured with an Agilent Infiniium DCA-J oscilloscope. The FWHM and jitter of the output signal are calculated using the oscilloscope. We assumed that the optical pulse was fast enough to present an impulse input and used fast Fourier transformation to calculate the frequency response and the device bandwidth. Note that this method underestimates the device bandwidth as it ignores the input pulse width. Pseudo-random binary sequence is generated using an Optellent OptoBERT 3200 generator (data rate ranges from 155 Mbps to 3.2 Gbps) to drive the 1.55-μm laser diode through a Mach–Zehnder modulator; the modulated optical signal is then coupled to the integrated T2-HPT. A tunable attenuator is used to change the coupled optical power. The output of the device is probed with a GSG probe and sent to an Agilent Infiniium DCA-J oscilloscope and a BERT analyzer to get the eye diagram and Q factor. This BERT system is limited to a maximum data rate of 3.2 Gbps. Capacitance–voltage (C–V) characteristics of HPTs with type-I and type-II band alignment are measured using an Agilent LCR meter HP 4285. The samples are prepared from the as-grown epitaxial structures on the InP substrate, and pillars with a diameter of 100 μm are etched and metallized. As shown in the inset of Fig. 1(b), n-doped InGaAs and InP constitute the collector (C) and emitter (E), respectively. P-doped InGaAs and p-doped GaAsSb constitute the base (B) for type-I and type-II, respectively; both are 50 nm thick and have a doping concentration of 5×10^{17} cm^{-3}.

C. Simulation

Three-dimensional FDTD simulation was performed using an FDTD tool (Lumerical), which was also used for calculating the transmission of light to the integrated detectors. The device simulation was performed using the ATLAS simulation software package. Simulations for integrating our T2-HPT detector with a 65 nm ASIC amplifier were conducted in Cadence Virtuoso 6.1.8 to evaluate the performance of the detector for applications as an optical receiver (detailed in the supplementary material).

SUPPLEMENTARY MATERIAL

See the supplementary material for the fabrication process, dark current, and −3 dB bandwidth of the HPT detectors with different sizes, thermal modeling of the integrated and as-grown detectors, energy consumption in the optical receiver system using photodetectors, comparison of the best reported on-chip optical receivers, and simulation of the T2-HPT integrated with ASIC amplifier.

ACKNOWLEDGMENTS

This work was partially supported by ARO Award No. W911NF1810429, NIH Award No. R21EY029516, and the W.M. Keck Foundation Award. The fabrication made use of the NUFAB facility of Northwestern University's NUANCE Center, the Pritzker Nanofabrication Facility part of the Pritzker School of Molecular Engineering at the University of Chicago, which have received support from the Soft and Hybrid Nanotechnology Experimental (SHyNE) Resource (Grant No. NSF ECCS-1542205); the MRSEC program (Grant No. NSF DMR-1720139) at the Materials Research Center; the International Institute for Nanotechnology (IIN); the Keck Foundation; and the State of Illinois, through the IIN, and the Center for Nanoscale Materials of Argonne National Laboratory. The use of the Center for Nanoscale Materials, an Office of Science user facility, was supported by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, under Contract No. DE-AC02-06CH11357. We thank Moshe Dolejsi (the University of Chicago) and David Czaplewski (Argonne National Laboratory) for their help in optimizing the EBL processes and material etching. We thank Alan Prosser (Fermilab) for his help in the BER measurement and the gigaBERT system.

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

L.L. did the designing, processing, measurements, and simulations of the integrated T2-HPT detectors. S.B. did the electron beam lithography (EBL) processing, the band energy simulation, and the speed modeling. S.W. put together the setup for high-speed measurement. L.L. and S.B. wrote the article. N.C. and F.F. did the circuit level simulation. H.M. conceived the idea, guided both experimental and modeling works and revised the article. All authors discussed the results and commented on the article.

Lining Liu: Data curation (equal); Formal analysis (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Simone Bianconi**: Data curation (equal); Formal analysis (equal); Visualization (equal); Writing – original draft (equal); Writing – review & editing (equal). **Skylar Wheaton**: Data curation (supporting); Writing – original draft (supporting). **Nathaniel Coirier**: Data curation (supporting); Visualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting). **Farah Fahim**: Data curation (supporting); Resources (supporting). **Hooman Mohseni**: Conceptualization (lead); Data curation (supporting); Formal analysis (supporting); Funding acquisition (lead); Investigation (lead); Methodology (equal); Project administration (lead); Supervision (lead); Validation (equal); Visualization (supporting); Writing – original draft (supporting); Writing – review & editing (supporting).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] S. Pai, Z. Sun, T. W. Hughes, T. Park, B. Bartlett, I. A. Williamson, M. Minkov, M. Milanizadeh, N. Abebe, F. Morichetti *et al.*, "Experimentally realized in situ backpropagation for deep learning in photonic neural networks," Science **380**, 398–404 (2023).

[2] J. Feldmann, N. Youngblood, M. Karpov, H. Gehring, X. Li, M. Stappers, M. Le Gallo, X. Fu, A. Lukashchuk, A. S. Raja *et al.*, "Parallel convolutional processing using an integrated photonic tensor core," Nature **589**, 52–58 (2021).

[3] W. Liu, M. Li, R. S. Guzzon, E. J. Norberg, J. S. Parker, M. Lu, L. A. Coldren, and J. Yao, "A fully reconfigurable photonic integrated signal processor," Nat. Photonics **10**, 190–195 (2016).

[4] J. Lee, A. Ingle, J. V. Chacko, K. W. Eliceiri, and M. Gupta, "CASPI: Collaborative photon processing for active single-photon imaging," Nat. Commun. **14**, 3158 (2023).

[5] J. Michel, J. Liu, and L. C. Kimerling, "High-performance Ge-on-Si photodetectors," Nat. Photonics **4**, 527–534 (2010).

[6] H. Shin, J. A. Cox, R. Jarecki, A. Starbuck, Z. Wang, and P. T. Rakich, "Control of coherent information via on-chip photonic–phononic emitter–receivers," Nat. Commun. **6**, 6427 (2015).

[7] D. A. B. Miller, "Physical reasons for optical interconnection," Int. J. Optoelectron. **11**, 155–168 (1997).

[8] J.-H. Han, F. Boeuf, J. Fujikata, S. Takahashi, S. Takagi, and M. Takenaka, "Efficient low-loss InGaAsP/Si hybrid MOS optical modulator," Nat. Photonics **11**, 486–490 (2017).

[9] L. Virot, P. Crozat, J.-M. Fédéli, J.-M. Hartmann, D. Marris-Morini, E. Cassan, F. Boeuf, and L. Vivien, "Germanium avalanche receiver for low power interconnects," Nat. Commun. **5**, 4957 (2014).

[10] G. Crosnier, D. Sanchez, S. Bouchoule, P. Monnier, G. Beaudoin, I. Sagnes, R. Raj, and F. Raineri, "Hybrid indium phosphide-on-silicon nanolaser diode," Nat. Photonics **11**, 297–300 (2017).

[11] D. A. Miller, "Device requirements for optical interconnects to silicon chips," Proc. IEEE **97**, 1166–1185 (2009).

[12] D. A. Miller, "Attojoule optoelectronics for low-energy information processing and communications," J. Lightwave Technol. **35**, 346–396 (2017).

[13] S. Rumley, D. Nikolova, R. Hendry, Q. Li, D. Calhoun, and K. Bergman, "Silicon photonics for exascale systems," J. Lightwave Technol. **33**, 547–562 (2015).

[14] M. Bahadori, S. Rumley, R. Polster, A. Gazman, M. Traverso, M. Webster, K. Patel, and K. Bergman, "Energy-performance optimized design of silicon photonic interconnection networks for high-performance computing," in *Design, Automation & Test in Europe Conference & Exhibition (DATE), 2017* (IEEE, 2017), pp. 326–331.

[15] C. Gasser, B. Goll, and H. Zimmermann, "Highly sensitive hybrid optical receiver with a compact 1.4 GHz linear transimpedance amplifier in 55 nm CMOS," Opt. Eng. **62**, 095103 (2023).

[16] V. Crozatier, M. Van Osta, R. Broucke, L. Bogaert, J. Van Kerrebrouck, N. Singh, and G. Torfs, "Experimental demonstration of a 25 GHz optoelectronic oscillator using an integrated resonant photoreceiver consisting of a GeSi photodiode and Co-designed III-V low-noise amplifier," 2024 International Topical Meeting on Microwave Photonics (MWP) 1–4 (2024).

[17] H. Chen, J. Verbist, P. Verheyen, P. De Heyn, G. Lepage, J. De Coster, P. Absil, X. Yin, J. Bauwelinck, J. Van Campenhout, and G. Roelkens, "High sensitivity 10 Gb/s Si photonic receiver based on a low-voltage waveguide-coupled Ge avalanche photodetector," Opt. Express **23**, 815–822 (2015).

[18] S. D. March, A. H. Jones, J. C. Campbell, and S. R. Bank, "Multistep staircase avalanche photodiodes with extremely low noise and deterministic amplification," Nat. Photonics **15**, 468–474 (2021).

[19] G. Williams, F. Capasso, and W. Tsang, "The graded bandgap multilayer avalanche photodiode: A new low-noise detector," IEEE Electron Device Lett. **3**, 71–73 (1982).

[20] S. Chandrasekhar, L. M. Lunardi, A. Gnauck, R. Hamm, and G. Qua, "High-speed monolithic p-i-n/HBT and HPT/HBT photoreceivers implemented with simple phototransistor structure," IEEE Photonics Technol. Lett. **5**, 1316–1318 (1993).

[21] D. Huber, R. Bauknecht, C. Bergamaschi, M. Bitter, A. Huber, T. Morf, A. Neiger, M. Rohner, I. Schnyder, V. Schwarz, and A. Jackel, "InP-InGaAs single HBT technology for photoreceiver OEIC's at 40 Gb/s and beyond," J. Lightwave Technol. **18**, 992 (2000).

[22] X. Hu, H. Zhang, D. Wu, D. Chen, L. Wang, X. Xiao, and S. Yu, "High-performance Germanium avalanche photodetector for 100 Gbit/s photonics receivers," Opt. Letters **46**, 3837–3840 (2021).

[23] D. Benedikovic, L. Virot, G. Aubin, J.-M. Hartmann, F. Amar, X. Le Roux, C. Alonso-Ramos, E. Cassan, D. Marris-Morini, P. Crozat *et al.*, "40 Gbps heterostructure germanium avalanche photo receiver on a silicon chip," Optica **7**, 775–783 (2020).

[24] X. Zeng, Z. Huang, B. Wang, D. Liang, M. Fiorentino, and R. G. Beausoleil, "Silicon–germanium avalanche photodiodes with direct control of electric field in charge multiplication region," Optica **6**, 772–777 (2019).

[25] S. Assefa, F. Xia, and Y. A. Vlasov, "Reinventing germanium avalanche photodetector for nanophotonic on-chip optical interconnects," Nature **464**, 80–84 (2010).

[26] Y. Kang, H.-D. Liu, M. Morse, M. J. Paniccia, M. Zadka, S. Litski, G. Sarid, A. Pauchard, Y.-H. Kuo, H.-W. Chen *et al.*, "Monolithic germanium/silicon avalanche photodiodes with 340 GHz gain–bandwidth product," Nat. Photonics **3**, 59–63 (2009).

[27] Y. Kang, Z. Huang, Y. Saado, J. Campbell, A. Pauchard, J. Bowers, and M. Paniccia, "High performance Ge/Si avalanche photodiodes development in intel," in *2011 Optical Fiber Communication Conference and Exposition and the National Fiber Optic Engineers Conference* (IEEE, 2011).

[28] J. Thuret, C. Gonzalez, J. Benchimol, M. Riet, and P. Berdaguer, "High-speed InP/InGaAs heterojunction phototransistor for millimetre-wave fibre radio communications," in *Conference Proceedings. Eleventh International Conference on Indium Phosphide and Related Materials (IPRM'99) (Cat. No.99CH36362)* (IEEE, 1999), pp. 389–392.

[29] J. Campbell, C. Burrus, A. G. Dentai, and K. Ogawa, "Small-area high-speed InP/InGaAs phototransistor," Appl. Phys. Lett. **39**, 820–821 (1981).

[30] D. Fritzsche, E. Kuphal, and R. Aulbach, "Fast response InP/InGaAs heterojunction phototransistors," Electron. Lett. **17**, 178–180 (1981).

[31] S. Chandrasekhar, M. K. Hoppe, A. G. Dentai, C. H. Joyner, and G. J. Qua, "Demonstration of enhanced performance of an InP/InGaAs heterojunction phototransistor with a base terminal," IEEE Electron Device Lett. **12**, 550–552 (1991).

[32] M.-S. Park, M. Rezaei, I. Nia, R. Brown, S. Bianconi, C. L. Tan, and H. Mohseni, "InGaAs/InP quantum well infrared photodetector integrated on Si substrate by Mo/Au metal-assisted wafer bonding," Opt. Mater. Express **8**, 413–419 (2018).

[33] A. Roy, M. Azadmehr, B. Q. Ta, P. Häfliger, and K. E. Aasmundtveit, "Design and fabrication of CMOS microstructures to locally synthesize carbon nanotubes for gas sensing," Sensors **19**, 4340 (2019).

[34] J. Jiang, J. H. Chu, and K. Banerjee, "CMOS-Compatible Doped-Multilayer-Graphene Interconnects for Next-Generation VLSI," in *2018 IEEE International Electron Devices Meeting (IEDM)* (IEEE, 2018), pp. 34–35.

[35] M. Kwakernaak, W. Chan, B. Kharas, N. Maley, H. Mohseni, L. Yang, D. Capewell, V. Frantz, T. Mood, G. Pajer *et al.*, "Amorphous silicon waveguide components for monolithic integration with InGaAsP gain sections," Optoelectron. Mater. Devices **6352**, 635217 (2006).

[36] A. Bonakdar and H. Mohseni, "Impact of optical antennas on active optoelectronic devices," Nanoscale **6**, 10961–10974 (2014).

[37] L. Tang, S. E. Kocabas, S. Latif, A. K. Okyay, D.-S. Ly-Gagnon, K. C. Saraswat, and D. A. Miller, "Nanometre-scale germanium photodetector enhanced by a near-infrared dipole antenna," Nat. Photonics **2**, 226–229 (2008).

[38] A. Melikyan, L. Alloatti, A. Muslija, D. Hillerkuss, P. C. Schindler, J. Li, R. Palmer, D. Korn, S. Muehlbrandt, D. Van Thourhout *et al.*, "High-speed plasmonic phase modulators," Nat. Photonics **8**, 229–233 (2014).

[39] S. Bianconi, M. Rezaei, M.-S. Park, W. Huang, C. L. Tan, and H. Mohseni, "Engineering the gain-bandwidth product of phototransistor diodes," Appl. Phys. Lett. **115**, 051104 (2019).

[40] M. Rezaei, M.-S. Park, C. L. Tan, and H. Mohseni, "Sensitivity limit of nanoscale phototransistors," IEEE Electron Device Lett. **38**, 1051–1054 (2017).

[41] L. Liu, J. Rabinowitz, S. Bianconi, M.-S. Park, and H. Mohseni, "Highly sensitive SWIR detector array based on nanoscale phototransistors integrated on cmos readout," Appl. Phys. Lett. **117**, 191102 (2020).

[42]J. Bailbé, L. Andrieux, A. Cazarré, T. Camps, A. Marty, J. Tasselli, and H. Granier, "Theory and experiment of the temperature dependence of GaAlAs/GaAs HBTs characteristics for power amplifier applications," Solid-State Electron. **38**, 279–286 (1995).
[43]M. Hafizi, W. E. Stanchina, R. A. Metzger, P. A. Macdonald, and F. Williams, "Temperature dependence of DC and RF characteristics of AlInAs/GaInAs HBT's," IEEE Trans. Electron Devices **40**, 1583–1588 (1993).
[44]C. Bolognesi, M. Dvorak, P. Yeo, X. Xu, and S. Watkins, "InP/GaAsSb/InP double HBTs: A new alternative for InP-based DHBTs," IEEE Trans. Electron Devices **48**, 2631–2639 (2001).
[45]N. Duan, T.-Y. Liow, A. E.-J. Lim, L. Ding, and G. Lo, "310 GHz gain-bandwidth product Ge/Si avalanche photodetector for 1550 nm light detection," Opt. Express **20**, 11031–11036 (2012).
[46]Z. Huang, C. Li, D. Liang, K. Yu, C. Santori, M. Fiorentino, W. Sorin, S. Palermo, and R. G. Beausoleil, "25 Gbps low-voltage waveguide Si–Ge avalanche photodiode," Optica **3**, 793–798 (2016).
[47]N. J. Martinez, C. T. Derose, R. W. Brock, A. L. Starbuck, A. T. Pomerene, A. L. Lentine, D. C. Trotter, and P. S. Davids, "High performance waveguide-coupled Ge-on-Si linear mode avalanche photodiodes," Opt. Express **24**, 19072–19081 (2016).
[48]A. R. Hawkins, W. Wu, P. Abraham, K. Streubel, and J. E. Bowers, "High gain-bandwidth-product silicon heterointerface photodetector," Appl. Phys. Lett. **70**, 303–305 (1997).
[49]T. Ochiai, T. Akazawa, Y. Miyatake, K. Sumita, S. Ohno, S. Monfray, F. Boeuf, K. Toprasertpong, S. Takagi, and M. Takenaka, "Ultrahigh-responsivity waveguide-coupled optical power monitor for Si photonic circuits operating at near-infrared wavelengths," Nat. Commun. **13**, 7443 (2022).
[50]M. Geis, S. Spector, M. Grein, J. Yoon, D. Lennon, and T. Lyszczarz, "Silicon waveguide infrared photodiodes with >35 GHz bandwidth and phototransistors with 50 AW-1 response," Opt. Express **17**, 5193–5204 (2009).
[51]A. Dehzangi, R. McClintock, D. Wu, A. Haddadi, R. Chevallier, and M. Razeghi, "Extended short wavelength infrared heterojunction phototransistors based on type II superlattices," Appl. Phys. Lett. **114**, 191109 (2019).
[52]Z. Xie, Z. Deng, J. Huang, Z. Xie, Z. Zhou, and B. Chen, "InP-based extended-short wave infrared heterojunction phototransistor," J. Lightwave Technol. **39**, 4814–4819 (2021).
[53]T. S. Kim, S.-H. Jeon, K. Ko, D.-H. Ahn, J.-H. Han, W. J. Choi, and K. J. Yu, "Fast, energy-efficient InGaAs synaptic phototransistors on flexible substrate," Adv. Electron. Mater. **9**, 2300437 (2023).
[54]Y. Li, A. Alian, M. Sivan, L. Huang, K. W. Ang, D. Lin, D. Mocuta, N. Collaert, and A. V.-Y. Thean, "A flexible InGaAs nanomembrane PhotoFET with tunable responsivities in near- and short-wave IR region for lightweight imaging applications," APL Mater. **7**, 031503 (2019).
[55]K. T. Settaluri, S. Lin, S. Moazeni, E. Timurdogan, C. Sun, M. Moresco, Z. Su, Y.-H. Chen, G. Leake, D. LaTulipe *et al.*, "Demonstration of an optical chip-to-chip link in a 3D integrated electronic-photonic platform," in *ESSCIRC Conference 2015 - 41st European Solid-State Circuits Conference (ESSCIRC)* (IEEE, 2015), pp. 156–159.
[56]G. H. Shin, C. Park, K. J. Lee, H. J. Jin, and S.-Y. Choi, "Ultrasensitive Phototransistor Based on WSe_2–MoS_2 van der Waals Heterojunction," Nano Lett. **20**, 5741–5748 (2020).