FERMILAB-PUB-23-672-ETD

Fast and Efficient Type-II Phototransistors Integrated on Silicon

Lining Liu[1], Simone Bianconi[1], Skylar Wheaton[1], Nathaniel Coirier[1], Farah Fahim[2], Hooman Mohseni[1]

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

Corresponding author: hmohseni@northwestern.edu;

arXiv:2311.03921v1 [physics.ins-det] 7 Nov 2023

Abstract

Increasing the efficiency and reducing the footprint of on-chip photodetectors enables dense optical interconnects for emerging computational and sensing applications. Avalanche photodetectors (APD) are currently the dominating on-chip photodetectors. However, the physics of avalanche multiplication leads to low energy efficiencies and prevents device operation at a high gain, due to a high excess noise, resulting in the need for electrical amplifiers. These properties significantly increase power consumption and footprint of current optical receivers. In contrast, heterojunction phototransistors (HPT) exhibit high efficiency and very small excess noise at high gain. However, HPT's gain-bandwidth product (GBP) is currently inferior to that of APDs at low optical powers. Here, we demonstrate that the type-II energy band alignment in an antimony-based HPT results in a significantly smaller junction capacitance and higher GBP at low optical powers. We used a CMOS-compatible heterogeneous integration method to create compact optical receivers on silicon with an energy efficiency that is about one order of magnitude higher than that of the best reported integrated APDs on silicon at a similar GBP of ∼270 GHz. Bitrate measurements show data rate spatial density above 800 Tbps/mm2, and an energy-per-bit consumption of only 6 fJ/bit at 3 Gbps. These unique features suggest new opportunities for creating highly efficient and compact on-chip optical receivers based on devices with type-II band alignment.

Keywords: optical interconnects, integrated optical receiver, fast photodetector, integrated photonics

1 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 interconnection [5–7]. Particularly, optical interconnects are considered to be one of the most promising candidates for next-generation on-chip communication [7] due 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

This manuscript has been authored by Fermi Research Alliance, LLC under Contract No. DE-AC02-07CH11359 with the U.S. Department of Energy, Office of Science, Office of High Energy Physics.

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 (see Supplementary, Figure S6). 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 CMOS amplifiers [15] require significant power and on-chip area (See Supplementary, Table SIII). Conversely, most state-of-the-art optical receivers such as APD and HPT 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 towards APD receivers in the past decade [9, 22–24]. However, beside 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, a high sensitivity, and a high energy efficiency. Our type-II heterojunction phototransistor (T2-HPT) integrated on silicon achieve a GBP similar to the best reported APDs on silicon, but with much lower dark current and operating bias voltage than APDs, 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 to 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 Methods). 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.

2 Results

2.1 Device Design

We used a low-temperature Mo/Au-Au/Mo bonding method [32] to transfer epitaxial layers to silicon substrates (See Supplementary I). We observed a typical root-mean-square (rms) roughness of $\sim$2 nm over large areas of the transferred film, indicating high-quality bonding (Fig. 1a-b). After the transfer, T2-HPTs with different sizes (ranging from 100 µm to 2 µm in diameter) were fabricated on the silicon substrate and integrated with a-Si:H waveguides and grating couplers (See Methods). a-Si:H can be easily deposited at low temperatures ($\leq$ 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 Methods and Supplementary I). Fig. 1c shows a scanning electron microscope (SEM) image of

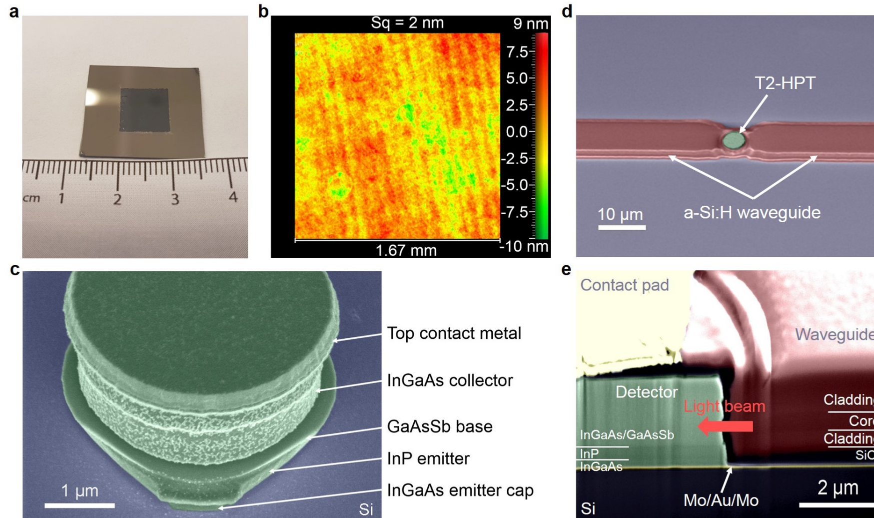

Fig. 1 Images of the integrated T2-HPT. **a**, Photograph of the InP/GaAsSb/InGaAs epitaxial layer transferred to silicon, and after the InP substrate was removed. **b**, Surface morphology of the transferred material shows rms roughness of about 2 nm. **c**, SEM image of the detector after pillar etching shows the epitaxial III-V semiconductor layers on the silicon substrate. **d**, SEM of the integrated T2-HPT after waveguide etching and before the BCB planarization. **e**, Cross-sectional FIB-SEM image of an integrated device at the end of the fabrication process.

a T2-HPT device before the waveguide integration, with clear collector, base and emitter layers of the HPT. The SEM images of a detector after waveguide integration show a top (Fig. 1d) and a cross-sectional (Fig. 1e) view of the interface between the waveguide and the detectors.

2.2 Device Measurement

We characterized the fully integrated T2-HPT devices with the architecture shown in Fig. 2a. While we measured devices with different diameters, here we mainly focus on the performance of devices of 2 µm diameter. We measured the dark current (I_d) and the photocurrent (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. 2b and Fig. 2c). Fig. 2d shows the dark and illuminated 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 II). The photocurrent shown in Fig. 2d was measured at an optical power of -19 dBm coupled to the 2 µm T2-HPT. Fig. 2e shows the DC responsivity at this power, exceeding 100 A/W at voltages as low as 0.75 V. 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 decreasing the width of the waveguides or introducing tapered waveguides. Furthermore, the low coupling to small devices can be addressed by employing hybrid optical antennae 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 Method). We coupled the output beam of a 1.55 µm laser diode, which was directly modulated by pseudorandom 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 BERT system— as shown in Fig. 3a. We measured sensitivities of -26 dBm and -25 dBm at 2.5 Gbps and 3.2 Gbps respectively, for a $BER < 10^9$ (Fig. 3b). More importantly, the high gain and direct driving capability of our device

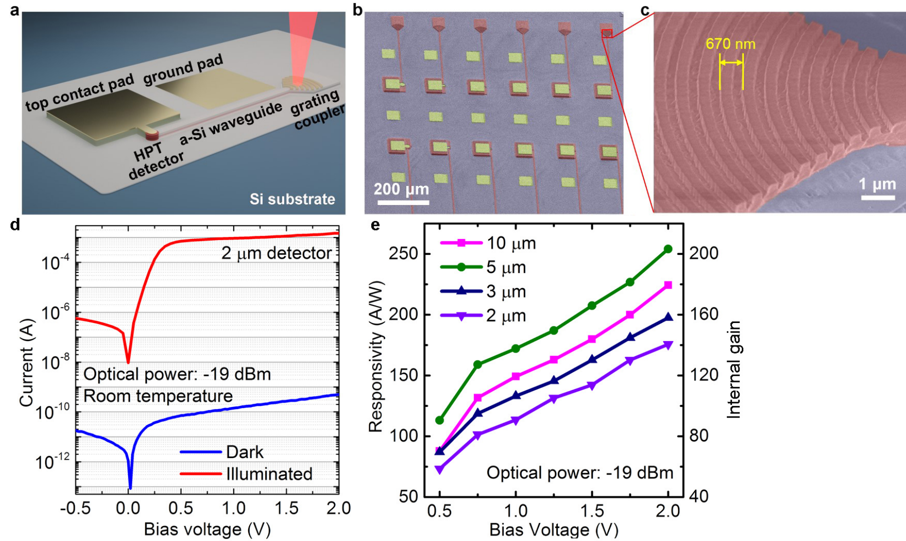

Fig. 2 Integrated T2-HPT detectors with top coupling. **a**, Schematic diagram of the measurement setup with the grating coupler. **b**, SEM image of the integrated detectors of different sizes. **c**, SEM image of the zoomed-in grating coupler with a period of 670 nm. **d**, Current-voltage characteristic of the integrated 2 µm detectors at room temperature at dark and under illumination. **e**, DC responsivity R of the detectors of different sizes as a function of bias voltage at an optical power of -19 dBm coupled to the detector.

eliminated the need for a high-gain and low-noise amplifier in these experiments and we directly coupled our device to the 50-Ohm load. Since our BERT system was limiting our measured data rate to 3.2 Gbps, we evaluated the actual bandwidth and the resulted estimated bit rate, using a fast optical pulse and a fast oscilloscope (see Methods). Fig. 3c shows the full width half maximum (FWHM) of the output signal of a 2 µm T2-HPT when receiving an optical pulse with FWHM of ∼7 ps. The electrical signal on a 50 Ohms load shows a FWHM of 97 ps and a jitter of less than 9 ps. The corresponding 3dB 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 section III of Supplementary). 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.

3 Discussion

It is known that type-II band alignment helps in preventing the onset of the base pushout 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 more sensitive, as we had proposed [40] and recently demonstrated experimentally [41]. Fig. 4 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

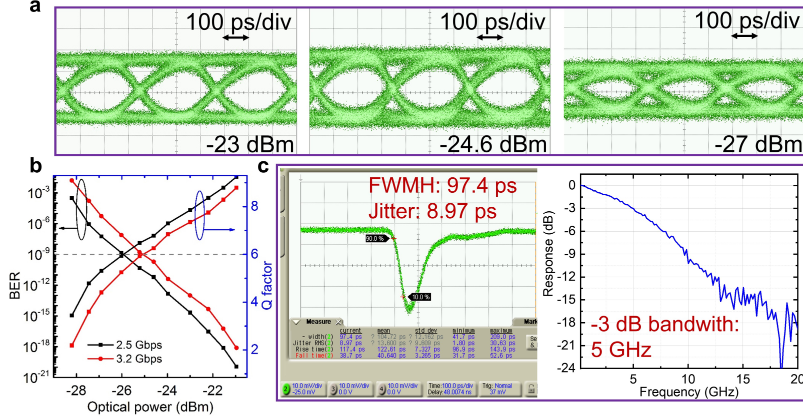

Fig. 3 Speed measurement of the integrated T2-HPT detectors. **a**, Eye diagrams of 2 μm device at 3.2 Gbps data rate, at different optical powers coupled to the detector. **b**, Bit error rate and Q factor as a function of optical power, extracted from the eye diagrams of the 2 μm detector. **c**, Output pulse from the integrated 2 μm detector and its Fourier transform.

level, thickness, and diameter. Our experimental results show that type-II HPT exhibits about six times smaller capacitance per unit area than type-I HPT at the bias voltage of 2 V utilized in this work (see Methods). The measured capacitance values and their bias dependencies are in good agreement with numerical simulations for both type-I and type-II devices (see Methods).

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 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 IV for thermal modeling). At an output power of 2 mW (2 V bias voltage and $/approx1$ mA photocurrent), the temperature increase in the transferred detector is more than 4 times lower than that in the as-grown detector. A high temperature can degrade the GBP of device, mainly due to the increased base transit time [42], as previously demonstrated for type-I and type-II HBT [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 and the energy waste per bit, and E_U is the energy delivered to the load per bit. For approaches based on optical receivers in CMOS, these values can be estimated as (See Supplementary V):

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

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

where C_L and V_L are the load capacitance and voltage (around 1 V in modern CMOS), I_d, I_{ph} and V_b are the dark current, photocurrent 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 Id due to the large electric field required for avalanche, both of which lead to poor energy efficiencies mostly around 1%. We compare the GBP versus 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]. Fig. 5a shows the GBP versus energy efficiency η and Fig. 5b shows the GBP versus consumed energy per bit E_C (details in Supplementary VI). Compared to integrated APDs, our integrated T2-HPT exhibits about one order of magnitude higher energy efficiency of $\eta \sim 15\%$, and a significantly lower energy

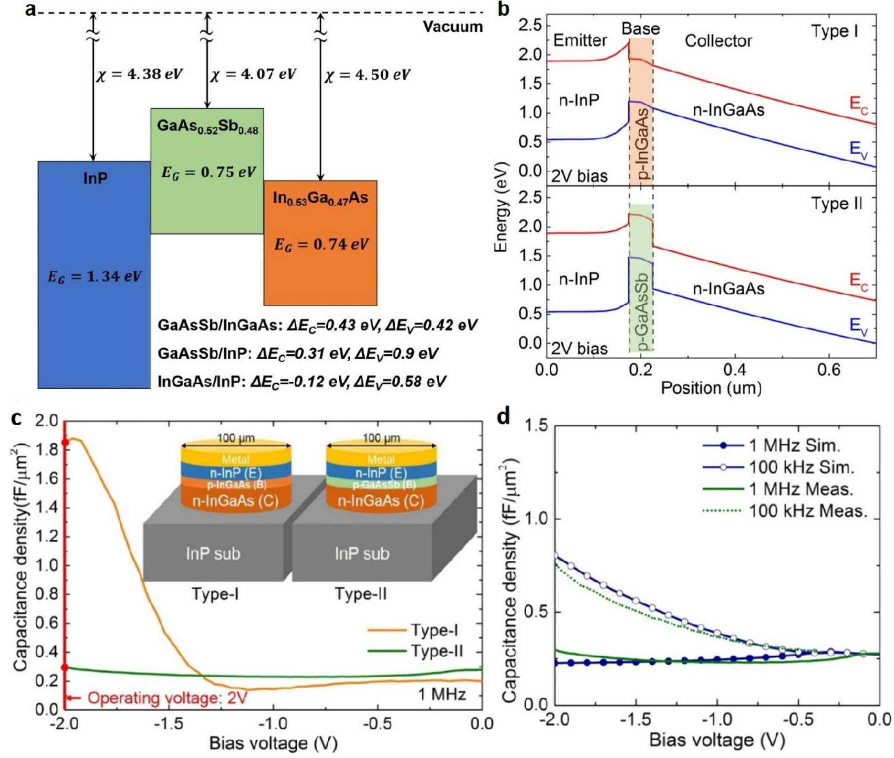

Fig. 4 Band structure of type I and type II. **a**, Schematic diagram of the energy band alignment of InGaAs, GaAsSb, and InP. GaAsSb forms a type-II alignment with InGaAs and InP, while InGaAs/InP forms a type-I alignment. **b**, Simulated band structures of two N-P-N HPTs based on type-I (top) and type-II (bottom) band alignments. **c**, Experimentally measured junction capacitance versus bias (C-V) for type-I and type-II HPTs with identical layer thickness and doping levels show that type-II HPT has a substantially lower junction capacitance at the operating bias of -2 Volt. **d**, Simulated C-V show good agreement with experimental data for low and high frequencies.

consumption, below $E_C \sim 6$ 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 photocurrent produced by the device, and hence by the sensitivity level and responsivity of the device. 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 X, and Table SII). 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 VII for the detail of the amplifier). The T2-HPT was modeled using the Ebers-Moll model, which showed a good agreement with the measured performance across different bias values and optical power (see Fig. S9a). Using the compact amplifier, the receiver could achieve a bitrate of 16.7 Gbps (Fig. S9b). When including capacitive loads of 5 fF or 50 fF, the highest open-eye data rates were ∼15 Gbps (Fig. S9c) and ∼10 Gbps (Fig. S9d) respectively. Crucially, the compact amplifier (Supplementary VII)

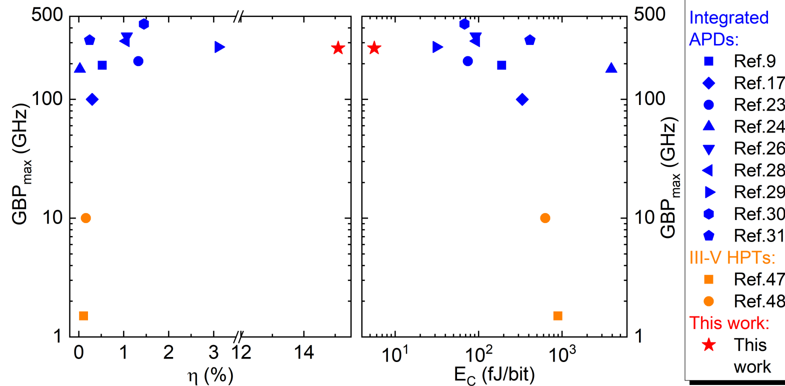

Fig. 5 Comparison of this work with the best reported detectors for optical receivers. **a**, GBPs vs. energy efficiency (η). Note that we removed a portion of the x-axis, from 2% to 12%, to increase the visibility. **b**, GBPs vs. total consumed energy (E_C) of detectors at internal gain of 15 and 2 fF load at optical power of about -30 dBm (see detailed calculation and tabulated data in Supplementary VI).

adds an energy dissipation of $\sim$ 60 fJ/bit, which is smaller than typical amplifiers used in other integrated receivers with a similar CMOS technology node, while achieving ~10 times higher sensitivity and about two times larger bitrate [49].

4 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 characterization 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 the best APDs. However, the T2-HPT showed about 10 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 ~6 fJ/bit. The unique combination of compactness, 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 in new material systems such as van der Waals heterostructures [50], which present an excellent opportunity for application of this strategy to a broad set of material and wavelengths.

Methods

Fabrication process

The T2-HPT detectors used in this work is grown by metalorganic chemical vapor deposition (MOCVD) on n-doped InP substrates. Fig. S3 (Supplementary Section II) shows the schematic diagrams of the process flow, which starts with the degreasing of a $\sim$ 1 cm × 1 cm sample and a $\sim$ 1.7 cm × 1.7 cm silicon substrate in organic solvents. Then, NH_4OH and $(NH_4)_2S$ treatment are 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 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, SiO2 (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 width of 5 µm and grating couplers with 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.

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 Calmar Optcom Femtosecond pulse laser as the light source emitting an optical pulse with a full width 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 by 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 our method underestimates the device bandwidth as it ignores the input pulse width. Pseudorandom binary sequence is generated by an Optellent OptoBERT 3200 (Data rate ranges from 155 Mbps to 3.2 Gbps) generator to drive the 1.55-µm laser diode through a Mach-Zehnder amplifier and 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. Note that our 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 by Agilent LCR meter HP 4285. The samples are prepared from the as-grown epitaxial structures on InP substrate, and pillars with diameter of 100 µm are etched and metallized. As shown in the inset of Fig. S4, 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 doping concentration of 5 × 1017 cm-3.

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 ATLAS simulation software package. Simulations for integrating our T2-HPT detector with a 65nm ASIC amplifier is conducted in Cadence Virtuoso 6.1.8 to evaluate the performance of the detector for applications as an optical receiver (detailed in Supplementary).

Acknowledgement

This work was partially supported by ARO award W911NF1810429, NIH award R21EY029516, and

the W.M. Keck Foundation Award. The fabrication made use of the NUFAB facility of Northwestern University's NUANCE Center, and of 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 (NSF ECCS-1542205); the MRSEC program (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 of the Center for Nanoscale Materials of Argonne National Laboratory. 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 would like to 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 would like to thank Alan Prosser (Fermilab) for his help in the BER measurement and the gigaBERT system.

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 manuscript. N.C. and F.F. did the circuit level simulation. H.M. conceived the idea, guided both experimental and modeling works and revised the manuscript. All authors discussed the results and commented on the manuscript.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request. Competing financial interest The authors declare no competing financial interest.

References

[1] Pai, S., Sun, Z., Hughes, T.W., Park, T., Bartlett, B., Williamson, I.A., Minkov, M., Milanizadeh, M., Abebe, N., Morichetti, F., *et al.*: Experimentally realized in situ backpropagation for deep learning in photonic neural networks. Science **380**(6643), 398–404 (2023)

[2] Feldmann, J., Youngblood, N., Karpov, M., Gehring, H., Li, X., Stappers, M., Le Gallo, M., Fu, X., Lukashchuk, A., Raja, A.S., *et al.*: Parallel convolutional processing using an integrated photonic tensor core. Nature **589**(7840), 52–58 (2021)

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

[4] Lee, J., Ingle, A., Chacko, J.V., Eliceiri, K.W., Gupta, M.: Caspi: collaborative photon processing for active single-photon imaging. Nature Communications **14**(1), 3158 (2023)

[5] Michel, J., Liu, J., Kimerling, L.C.: High-performance ge-on-si photodetectors. Nature photonics **4**(8), 527–534 (2010)

[6] Shin, H., Cox, J.A., Jarecki, R., Starbuck, A., Wang, Z., Rakich, P.T.: Control of coherent information via on-chip photonic–phononic emitter–receivers. Nature communications **6**(1), 6427 (2015)

[7] Miller, D.A.: Physical reasons for optical interconnection. Intel J. Optoelectronics **11**, 155–168 (1997)

[8] Han, J.-H., Boeuf, F., Fujikata, J., Takahashi, S., Takagi, S., Takenaka, M.: Efficient low-loss ingaasp/si hybrid mos optical modulator. Nature Photonics **11**(8), 486–490 (2017)

[9] Virot, L., Crozat, P., Fédéli, J.-M., Hartmann, J.-M., Marris-Morini, D., Cassan, E., Boeuf, F., Vivien, L.: Germanium avalanche receiver for low power interconnects. Nature communications **5**(1), 4957 (2014)

[10] Crosnier, G., Sanchez, D., Bouchoule, S.,

Monnier, P., Beaudoin, G., Sagnes, I., Raj, R., Raineri, F.: Hybrid indium phosphide-on-silicon nanolaser diode. Nature Photonics **11**(5), 297–300 (2017)

[11] Miller, D.A.: Device requirements for optical interconnects to silicon chips. Proceedings of the IEEE **97**(7), 1166–1185 (2009)

[12] Miller, D.A.: Attojoule optoelectronics for low-energy information processing and communications. Journal of Lightwave Technology **35**(3), 346–396 (2017)

[13] Rumley, S., Nikolova, D., Hendry, R., Li, Q., Calhoun, D., Bergman, K.: Silicon photonics for exascale systems. Journal of Lightwave Technology **33**(3), 547–562 (2015)

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

[15] Gasser, C., Goll, B., Zimmermann, H.: Highly sensitive hybrid optical receiver with a compact 1.4 ghz linear transimpedance amplifier in 55 nm cmos. Optical Engineering **62**(9), 095103–095103 (2023)

[16] Huang, Z., Li, C., Liang, D., Yu, K., Santori, C., Fiorentino, M., Sorin, W., Palermo, S., Beausoleil, R.G.: 25 gbps low-voltage waveguide si–ge avalanche photodiode. Optica **3**(8), 793–798 (2016)

[17] Chen, H., Verbist, J., Verheyen, P., De Heyn, P., Lepage, G., De Coster, J., Absil, P., Yin, X., Bauwelinck, J., Van Campenhout, J., *et al.*: High sensitivity 10gb/s si photonic receiver based on a low-voltage waveguide-coupled ge avalanche photodetector. Optics express **23**(2), 815–822 (2015)

[18] March, S.D., Jones, A.H., Campbell, J.C., Bank, S.R.: Multistep staircase avalanche photodiodes with extremely low noise and deterministic amplification. Nature Photonics **15**(6), 468–474 (2021)

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

[20] Chandrasekhar, S., Lunardi, L.M., Gnauck, A., Hamm, R., Qua, G.: High-speed monolithic pin/hbt and hpt/hbt photoreceivers implemented with simple phototransistor structure. IEEE photonics technology letters **5**(11), 1316–1318 (1993)

[21] Huber, D., Bauknecht, R., Bergamaschi, C., Bitter, M., Huber, A., Morf, T., Neiger, A., Rohner, M., Schnyder, I., Schwarz, V., *et al.*: Inp-ingaas single hbt technology for photoreceiver oeic's at 40 gb/s and beyond. Journal of lightwave technology **18**(7), 992 (2000)

[22] Chen, H., Verbist, J., Verheyen, P., De Heyn, P., Lepage, G., De Coster, J., Absil, P., Yin, X., Bauwelinck, J., Van Campenhout, J., *et al.*: High sensitivity 10gb/s si photonic receiver based on a low-voltage waveguide-coupled ge avalanche photodetector. Optics express **23**(2), 815–822 (2015)

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

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

[25] Assefa, S., Xia, F., Vlasov, Y.A.: Reinventing germanium avalanche photodetector for nanophotonic on-chip optical interconnects. Nature **464**(7285), 80–84 (2010)

[26] Kang, Y., Liu, H.-D., Morse, M., Paniccia, M.J., Zadka, M., Litski, S., Sarid,

G., Pauchard, A., Kuo, Y.-H., Chen, H.-W., *et al.*: Monolithic germanium/silicon avalanche photodiodes with 340 ghz gain–bandwidth product. Nature photonics **3**(1), 59–63 (2009)

[27] Kang, Y., Huang, Z., Saado, Y., Campbell, J., Pauchard, A., Bowers, J., Paniccia, M.: High performance ge/si avalanche photodiodes development in intel. In: Optical Fiber Communication Conference, p. 1 (2011). Optica Publishing Group

[28] Thuret, J., Gonzalez, C., Benchimol, J., Riet, M., Berdaguer, P.: 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), pp. 389–392 (1999). IEEE

[29] Campbell, J., Burrus, C., Dentai, A.G., Ogawa, K.: Small-area high-speed inp/ingaas phototransistor. Applied physics letters **39**(10), 820–821 (1981)

[30] Fritzsche, D., Kuphal, E., Aulbach, R.: Fast response inp/ingaasp heterojunction phototransistors. Electronics Letters **5**(17), 178–180 (1981)

[31] Chandrasekhar, S., Hoppe, M.K., Dentai, A.G., Joyner, C.H., Qua, G.J.: Demonstration of enhanced performance of an inp/ingaas heterojunction phototransistor with a base terminal. IEEE Electron Device Letters **12**(10), 550–552 (1991)

[32] Park, M.-S., Rezaei, M., Nia, I., Brown, R., Bianconi, S., Tan, C.L., Mohseni, H.: Ingaas/inp quantum well infrared photodetector integrated on si substrate by mo/au metal-assisted wafer bonding. Optical Materials Express **8**(2), 413–419 (2018)

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

[34] Jiang, J., Chu, J.H., Banerjee, K.: Cmos-compatible doped-multilayer-graphene interconnects for next-generation vlsi. In: 2018 IEEE International Electron Devices Meeting (IEDM), pp. 34–5 (2018). IEEE

[35] Kwakernaak, M., Chan, W., Kharas, B., Maley, N., Mohseni, H., Yang, L., Capewell, D., Frantz, V., Mood, T., Pajer, G., *et al.*: Amorphous silicon waveguide components for monolithic integration with ingaasp gain sections. In: Optoelectronic Materials and Devices, vol. 6352, pp. 292–296 (2006). SPIE

[36] Bonakdar, A., Mohseni, H.: Impact of optical antennas on active optoelectronic devices. Nanoscale **6**(19), 10961–10974 (2014)

[37] Tang, L., Kocabas, S.E., Latif, S., Okyay, A.K., Ly-Gagnon, D.-S., Saraswat, K.C., Miller, D.A.: Nanometre-scale germanium photodetector enhanced by a near-infrared dipole antenna. Nature Photonics **2**(4), 226–229 (2008)

[38] Melikyan, A., Alloatti, L., Muslija, A., Hillerkuss, D., Schindler, P.C., Li, J., Palmer, R., Korn, D., Muehlbrandt, S., Van Thourhout, D., *et al.*: High-speed plasmonic phase modulators. Nature Photonics **8**(3), 229–233 (2014)

[39] Bianconi, S., Rezaei, M., Park, M.-S., Huang, W., Tan, C.L., Mohseni, H.: Engineering the gain-bandwidth product of phototransistor diodes. Applied physics letters **115**(5) (2019)

[40] Rezaei, M., Park, M.-S., Tan, C.L., Mohseni, H.: Sensitivity limit of nanoscale phototransistors. IEEE Electron Device Letters **38**(8), 1051–1054 (2017)

[41] Liu, L., Rabinowitz, J., Bianconi, S., Park, M.-S., Mohseni, H.: Highly sensitive swir detector array based on nanoscale phototransistors integrated on cmos readout. Applied Physics Letters **117**(19) (2020)

[42] Bailbe, J., Andrieux, L., Cazarré, A., Camps, T., Marty, A., Tasselli, J., Granier, H.: Theory and experiment of the temperature

dependence of gaalas/gaas hbts characteristics for power amplifier applications. Solid-state electronics **38**(2), 279–286 (1995)

[43] Hafizi, M., Stanchina, W.E., Metzger, R.A., Macdonald, P.A., Williams, F.: Temperature dependence of dc and rf characteristics of alinas/gainas hbt's. IEEE transactions on electron devices **40**(9), 1583–1588 (1993)

[44] Bolognesi, C., Dvorak, M., Yeo, P., Xu, X., Watkins, S.: Inp/gaassb/inp double hbts: a new alternative for inp-based dhbts. IEEE Transactions on Electron Devices **48**(11), 2631–2639 (2001)

[45] Duan, N., Liow, T.-Y., Lim, A.E.-J., Ding, L., Lo, G.: 310 ghz gain-bandwidth product ge/si avalanche photodetector for 1550 nm light detection. Optics express **20**(10), 11031–11036 (2012)

[46] Huang, Z., Li, C., Liang, D., Yu, K., Santori, C., Fiorentino, M., Sorin, W., Palermo, S., Beausoleil, R.G.: 25 gbps low-voltage waveguide si–ge avalanche photodiode. Optica **3**(8), 793–798 (2016)

[47] Martinez, N.J., Derose, C.T., Brock, R.W., Starbuck, A.L., Pomerene, A.T., Lentine, A.L., Trotter, D.C., Davids, P.S.: High performance waveguide-coupled ge-on-si linear mode avalanche photodiodes. Optics express **24**(17), 19072–19081 (2016)

[48] Hawkins, A.R., Wu, W., Abraham, P., Streubel, K., Bowers, J.E.: High gain-bandwidth-product silicon heterointerface photodetector. Applied Physics Letters **70**(3), 303–305 (1997)

[49] Settaluri, K.T., Lin, S., Moazeni, S., Timurdogan, E., Sun, C., Moresco, M., Su, Z., Chen, Y.-H., Leake, G., LaTulipe, D., *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), pp. 156–159 (2015). IEEE

[50] Shin, G.H., Park, C., Lee, K.J., Jin, H.J., Choi, S.-Y.: Ultrasensitive phototransistor based on wse2–mos2 van der waals heterojunction. Nano Letters **20**(8), 5741–5748 (2020)

arXiv:2311.03921v1 [physics.ins-det] 7 Nov 2023

Supplementary Information: Fast and Efficient Type-II Phototransistors Integrated on Silicon

Lining Liu[1], Simone Bianconi[1], Skylar Wheaton[1], Nathaniel Coirier[1], Farah Fahim[2], Hooman Mohseni[1]

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

Corresponding author: hmohseni@northwestern.edu;

Contents

1 Thin-film transfer procedure and fabrication process **2**

2 Dark current of detectors with different sizes **3**

3 Model for -3 dB bandwidth of detectors with different sizes **3**

4 Thermal modeling of the integrated and as-grown detector **5**

5 Energy consumption in optical receiver system using photodetectors **7**

6 Comparison of the best reported on-chip optical receivers **9**

7 Simulation of T2-HPT integrated with an ASIC amplifier **11**

1 Thin-film transfer procedure and fabrication process

The epitaxial structure shown in Table 1 was grown on InP substrates. Large-area (∼ 1 cm × 1 cm) epitaxial layers were successfully transferred to silicon substrates by metal-assisted wafer bonding and subsequent wet etching of the InP substrate. Au-Au bonding has been shown to provide a shear bonding strength as high as 20 MPa under a low-temperature process that prevents residual stress responsible for bowing or cracks [1–3]. Mo has a high thermal conductivity (139 W/mK), an excellent contact resistance for III-V semiconductors, and can be used as diffusion barrier to prevent the unfavorable diffusion of Au atoms [4]. Therefore, Mo/Au is chosen as the metal stack to assist the wafer bonding. The mirror-like surface of the transferred film is crack-free and continuous as shown in Fig. 1a in the manuscript. This corroborates the high-quality of the wafer-bonding and chemical InP substrate removal.

Material	Doping	Thickness
$In_{0.53}Ga_{0.47}As$	N-type, $10^{19} cm^{-3}$	100 nm
InP	N-type, 10^{17} to $10^{19} cm^{-3}$ graded	50 nm
InP	N-type, $10^{17} cm^{-3}$	200 nm
$GaAs_{0.52}Sb_{0.48}$	P-type, $5x10^{17} cm^{-3}$	50 nm
$In_{0.53}Ga_{0.47}As$	N-type, $10^{15} cm^{-3}$	1000 nm
$In_{0.53}Ga_{0.47}As$	N-type, 10^{15} to $10^{18} cm^{-3}$ graded	50 nm
InP substrate		

Table 1 Epitaxial structure of the HPT wafer grown on InP substrates by low-pressure metal-organic chemical vapor deposition (LP-MOCVD).

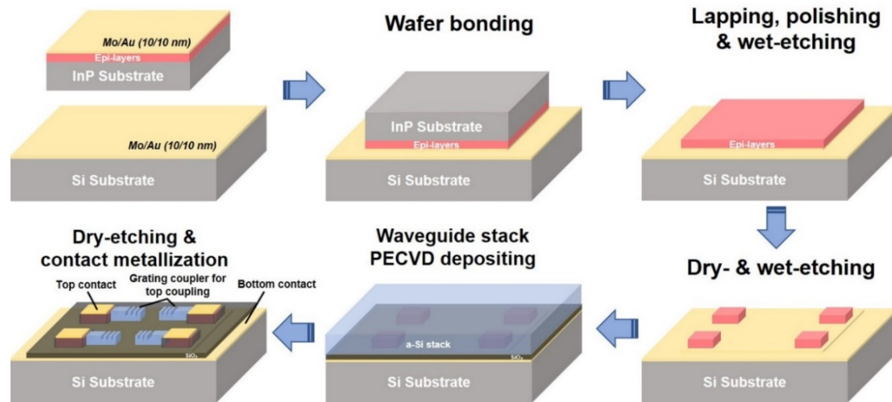

Fig. 1 Fabrication process of the integrated T2-HPT detector.

Table 1 shows the epitaxial structure of the HPT device in this work which uses Sb-based type-II band alignment. The fabrication process of the integrated detectors,

waveguides and grating couplers is depicted in Fig. 1 and described in the Method section in the manuscript. The a-Si:H waveguides and grating couplers consist of a cladding/core/cladding layer stack with a refractive index profile of 2.4/2.8/2.4 at 1.55 µm. a-Si:H is ideal for the optical waveguides and couplers used in this work, due to its high refractive index and low optical loss. It can be easily deposited by plasma-enhanced chemical vapor deposition (PECVD) at low temperatures (*leq* 300 °C). Besides, its refractive index can be controlled by adjusting the RF power and chamber pressure during deposition.

2 Dark current of detectors with different sizes

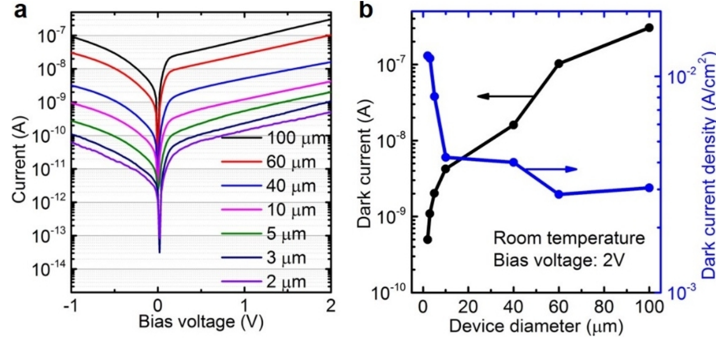

Fig. 2 Dark current. **a**, Dark current versus bias voltage for detectors with different sizes. **b**, Dark current and dark current density at 2 V bias extracted from **a**.

Fig. 2a shows the dark current of detectors with different sizes at room temperature. The detectors with diameter smaller than 10 µm exhibit dark current in nA level at 2 V, which is promising for low energy applications. The dark current density (Fig. 2b) increases with decreasing detector size, especially under 10 µm. This is likely caused by the surface effect playing a more dominant role in devices of smaller dimension [5].

3 Model for -3 dB bandwidth of detectors with different sizes

The -3 dB bandwidth of a device is related to its response time as $BW_{3dB} \approx 0.35/\tau_r$ [6]. It has been shown that the rise time of HPTs can be effectively approximated by [7]:

$$\tau_r = 2.2(\tau_{R_L} + r_d C_T + RC_L) \quad (1)$$

where τ_{R_L} is the recombination lifetime in the base, r_d is the transistor's dynamic resistance, C_T is the total junction capacitance and the last term represents the RC time constant of the load driven by the detector. By comparing this equation to the

measured response time and -3dB bandwidth for devices of different sizes, ranging from 2 µm to 100 µm in diameter, it is possible to separate the contribution of the device RC time constant in the equation above, since it is the only term that depends on the device area. The recombination lifetime is assumed constant for all devices, since the recombination sites are mainly localized at the interface between the base and the emitter layers [8]. Upon shrinking the size of the devices, surface states can provide additional recombination sites, which could explain the slight decrease in response time for the 2 µm device; however, this effect was not noticed in any of devices larger than 2 µm in diameter. The junction capacitance of the devices was approximated by a plate parallel capacitor, as $C = (\varepsilon_0 \varepsilon_r A)/w$, where ε_0 and ε_r are the vacuum permittivity and the semiconductor dielectric constant respectively, and A is the cross sectional area of the devices. The depletion width w of the devices was estimated using a commercial numerical simulation software package (ATLAS from Silvaco International). The dynamic resistance was calculated as [9]:

$$r_d = \frac{\eta_F kT}{q} \frac{1}{i_D + I_{ph}} \tag{2}$$

where eta_F is the ideality factor, k is Boltzmann's constant, T is temperature, q is the electron charge, i_D and i_{ph} are the devices internal dark and photo-current, respectively. The ideality factor was calculated based on the optical power level, using the expression proposed by Movassaghi et al. [10] for a class of very similar devices, yielding $\eta_F \approx 2.6$.

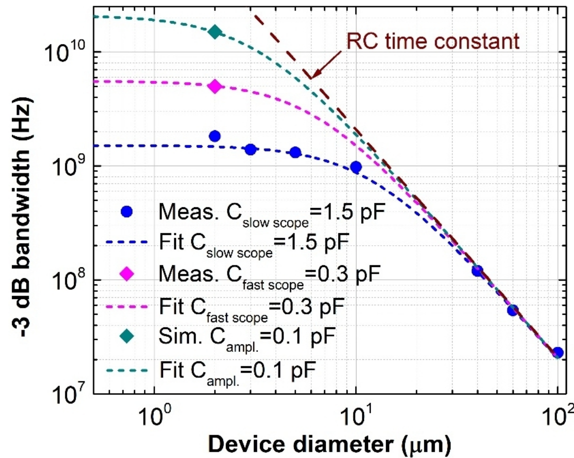

Fig. 3 Photoresponse bandwidth model of detectors with different sizes and its fitting to measurement data.

The results of the model are represented by the dashed curves in Fig. 3, for three different values of the load capacitance, corresponding to different measurements and simulations. In this plot, the device RC time constant as a function of its diameter is represented by the brown dashed line. Larger devices are characterized by larger

Device	GBP max (GHz)	Minimum measured rise time (ns)	Integrated on Si
Previous Device#1 [11]	1.2	32.5	No
Previous Device#2 [11]	4.9	20	No
Previous Device#3[11]	0.9	395	No
Previous Device#4 [12]	6	5	No
This work	270	0.04	Yes

Table 2 Comparison of this work with our previous published works based on HPT

capacitance, and are therefore limited in bandwidth by their RC time constant. Scaling the devices to small sizes allows to decrease the RC contribution due to their response time, and their bandwidth becomes ultimately limited by the recombination lifetime and the capacitance of the load. The smallest devices we measured (2 µm) are limited in bandwidth by the load capacitance of the measurement instrumentation. Our experimental results show reducing the capacitance of the oscilloscope results in an improvement in bandwidth, as shown in Fig. 3. In addition, integrating the devices with a low-capacitance ASIC amplifier could further increase their -3dB bandwidth, as shown by our detailed simulation using Cadence Virtuoso, and represented by the green diamond in Fig. 3. Note that the high internal gain of the T2-HPT allows a very compact and simple amplifier design, with very low power consumption and small footprint. The details of the simulation and ASIC design are given in the following section.

4 Thermal modeling of the integrated and as-grown detector

Table 2 compares the GBP and speed of devices we previously reported to those achieved in this work, which are significantly superior [11, 12]. While in this work we have fabricated devices smaller than ever before, we note that size scaling is not the only factor improving the performance of the detector. For example, Device #2 in Table 2 has the same epitaxial structure as the ones used in this work, however, when comparing detectors of identical size of 30 µm, the GBP in this work is about 6 time higher. The devices in our previous reports present significant differences compared with the ones presented in this work. Crucially, in our previous work, the detectors were fabricated on the InP native substrate (as-grown) and back-illuminated, while in this work the epitaxial layer is transferred to a silicon substrate and subsequently fabricated on silicon, and side-illuminated through integrated a-Si waveguides.

To conclusively compare the integrated and as-grown detectors, we made a sample with 2 µm devices on native InP substrate (as-grown), using the exact same epitaxial material. As expected, the DC performance of the two sets of devices are almost identical. Notably, the dark current of the transferred and as-grown detectors are very similar (Fig. 4a), indicating that transferring the detector to Si substrate has no significant effect on the dark current of the detector. In order to rule-out any possible differences or limitations in the measurement setup, we conducted a femtosecond laser pulse measurements using the same measurement setup both for the transferred and

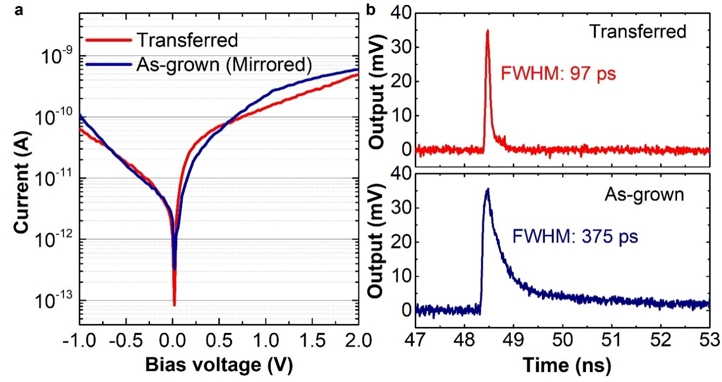

Fig. 4 Comparison of transferred and as-grown HPT detectors. Comparisons of **a**, dark current and **b**, output pulses from the transferred and as-grown T2-HPTs. Note that after bonded to Si substrate, the transferred detector is upside down with respect to the as-grown detector, so its bias voltage should be reversed with each other. For better comparison, the dark current curve of the as-grown detector is mirrored in panel **a**.

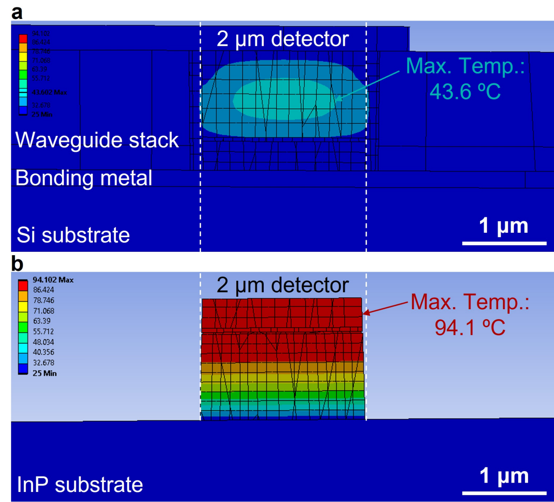

Fig. 5 Comparison of thermal conductance between the as-grown and integrated detectors. Thermal simulation of the **a**, transferred and **b**, as-grown detectors.

as-grown detectors. When comparing the speed of both devices, we observed that the transferred detector is ~ 4 times faster than the as-grown detector at the same output amplitude, as shown in Fig. 4b. Therefore, with exactly the same size and measurement conditions, the transferred detectors conclusively exhibit higher speed the as-grown detectors. The exact physics behind the marked difference is not clear at this time and will be a subject of our future work. However, we hypothesize that it is related to the significantly higher thermal and electrical conductivity of the back

metal contact when devices are transferred to Si substrate, and the lower diffusion length needed when the device is side illuminated rather than back-side illuminated.

In support of this hypothesis, we performed a detailed 3D thermal simulation using ANSYS thermal modeling (FEM) tool to compare the as-grown and integrated detectors. The power generated from the photo current is assumed to be 2 mW, since the photocurrent in Fig. 2d in the manuscript is in mA level and the operating bias voltage is 2 V. Assuming the temperature in the surrounding areas is 25 °C, the temperature increase during detection in the transferred detector ($\Delta T = 18.6°C$) is significantly lower than that of the as-grown detector ($\Delta T = 69.1°C$), as shown in Fig. 5.

5 Energy consumption in optical receiver system using photodetectors

To better understand the sources of energy waste in an optical receiver, let us first examine the energy flow. In an optical receiver, data is transduced from the optical domain to the electrical domain (Fig. 6). For most practical applications, the energy of the signal must also be boosted significantly since the input optical energy is typically less than the electrical energy required to drive the capacitive load of the connected electronics. This is commonly achieved using p-i-n or avalanche photodiodes (APD) combined with an electrical amplifier.

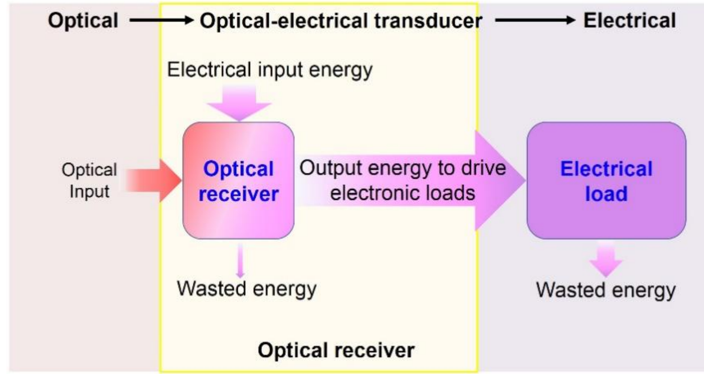

Fig. 6 Energy flow of the system for optical interconnects. The optical receiver transduces the input optical signal to an electrical signal for on-chip data transmission. The input energy of the system consists of an optical signal and an electrical bias, part of which is wasted by the dark current of the optical receiver and the rest of which is transduced to the electrical load. In the electrical load stage, part of the energy is utilized while the other part is wasted. Calculation of these energies is discussed in the text.

Here we consider the simplest general equivalent circuit that could be used for both HPT and APD detectors to create an optical receiver system. Since the detector always needs a bias Vb, in order to keep the lower side of the load at ground, a negative voltage bias is applied (-V_b). A cascode transistor is the simplest circuit to

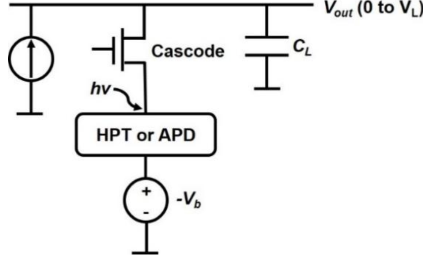

Fig. 7 Equivalent circuit for an optical receiver based on a photodetector.

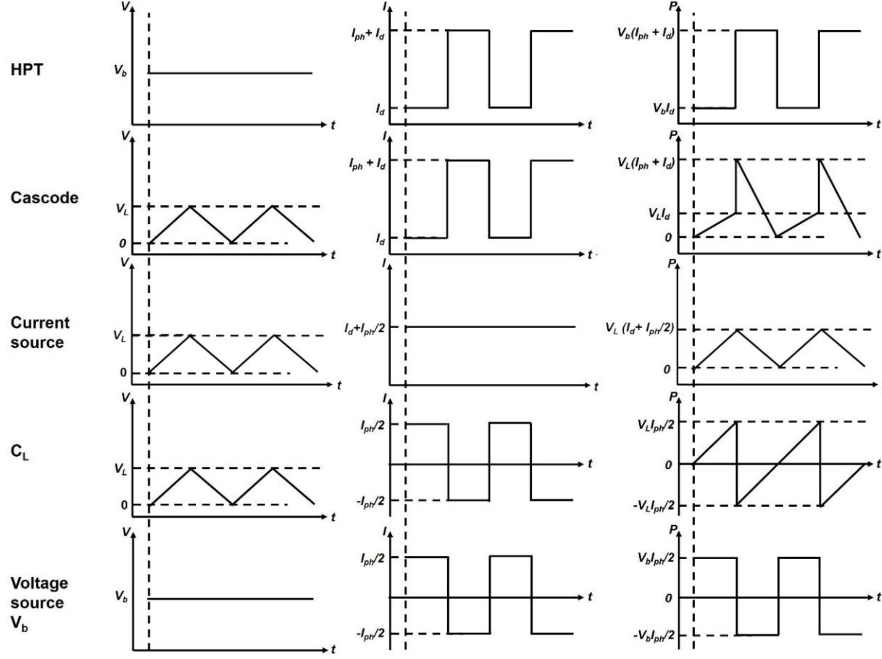

Fig. 8 Current, voltage and power consumption of each element in the circuit.

stabilize the detector bias around Vb, which is crucial to maintaining the gain and speed of HPTs and APDs while transferring its current to the load. Here, we assumed an ideal transistor and current source, although these assumption have little effect on the general conclusions that can be made from this model (Fig. 7).

The energy dissipation and efficiency can be calculated from the current and voltage of each element in this circuit. Here we used the conventional approach of assuming an equal probability for the "0" and "1" bits (alternative modulation schemes with unbalanced probabilities result in a similar overall energy conclusion). The first and second column in Fig. 8 show the current and voltage of each element respectively. The instantaneous power dissipated by each element (third column) can be obtained from the product of its current and voltage. Therefore, the energy dissipated per cycle

	State-of-the-art PIN + TIA	Type-II HPT in this work
Area	410 µm × 410 µm	2 µm in diameter
Power consumption	31 mW (TIA)	2 mW (no amplifier)
Data rate with high quality eye diagram with lowest optical power	2 Gbps with -27.45 dBm	3.2 Gbps with -27 dBm
3 dB bandwidth	1.4 GHz	5 GHz
Waveguide-coupled	No	Yes

Table 3 Comparison between one of the best reported p-i-n optical receiver with external amplifier and the receiver presented in this work.

for each element (denoted as $E_{detector}$, $E_{cascode}$, E_{cs}, E_{load}, and E_{vs} for the detector, cascode transistor, current source, load and voltage source, respectively) equals to the integration of its instantaneous power over time (Eq. 3 to 6).

$$E_{detector} = \frac{1}{2}\frac{V_b I_{ph} + 2V_b I_d}{BR} \tag{3}$$

$$E_{cascode} = \frac{1}{2}\frac{V_L I_{ph}/2 + V_L I_d}{BR} \tag{4}$$

$$E_{cs} = \frac{1}{2}\frac{V_L I_{ph}/2 + V_L I_d}{BR} \tag{5}$$

$$E_{load} = E_{vs} = 0 \tag{6}$$

And since the energy utilized by the load E_U is:

$$E_U \sim \frac{1}{2}C_L V_L^2 \tag{7}$$

The receiver energy consumption is:

$$E_C = E_{detector} + E_{cascode} + E_{cs} = \frac{2V_b I_d + V_b I_{ph} + 2V_L I_d + V_L I_{ph}}{2BR} \tag{8}$$

Therefore, the energy efficiency is:

$$\eta \sim \frac{E_U}{E_C + E_U} = \frac{\frac{1}{2}C_L V_L^2}{\frac{2V_b I_d + V_b I_{ph} + 2V_L I_d + V_L I_{ph}}{2BR} + \frac{1}{2}C_L V_L^2} \tag{9}$$

6 Comparison of the best reported on-chip optical receivers

First we consider the case of photodetectors without internal gain (e.g. p-i-n photodiode), which require external amplifiers. Table 3 compares one of the best reported optical receiver based on p-i-n photodiode integrated with a custom-designed CMOS transimpedance amplifier (TIA) [13] with the optical receiver presented in this work, operating without an external amplifier. The latter shows higher bandwidth, much smaller area and one order of magnitude lower power consumption, at a similar sensitivity level.

Ref.	Type	GBP_{max} (GHz)	R (A/W)	V_b (V)	I_d (A)	I_{ph} (A)	BR_{max} (Gbps)	E_C (fJ/bit)	η (%)
5	Ge APD	193	6	6.3	$3.6\ 10^{-5}$	$6.0\ 10^{-6}$	1.5	190	0.52
17	Ge/Si APD	100	9	6.2	$1.0\ 10^{-4}$	9.010^{-6}	2.3	334	0.30
23	Ge APD	210	7.35	9	$1.0\ 10^{-5}$	$7.4\ 10^{-6}$	1.9	74.4	1.33
24	Ge/Si APD	180	7.2	6	$1.0\ 10^{-3}$	$7.2\ 10^{-6}$	1.8	3900	0.03
26	Ge/Si APD	340	8.25	24	$3.5\ 10^{-6}$	$8.2\ 10^{-6}$	2.1	92.4	1.07
28	Ge/Si APD	310	4.5	26.5	$1.6\ 10^{-6}$	$4.5\ 10^{-6}$	1.1	94.1	1.05
29	Ge/Si APD	276	15.75	9.8	$3.5\ 10^{-6}$	$1.6\ 10^{-5}$	3.9	31.2	3.11
30	Ge/Si APD	432	12	28.9	$8.1\ 10^{-7}$	$1.2\ 10^{-5}$	3.0	67.9	1.45
31	InGaAs APD	315	8.55	21.5	$3.5\ 10^{-5}$	$8.6\ 10^{-6}$	2.1	413	0.24
45	InP/InGaAs HPT	62	NA	1.6	$1.0\ 10^{-7}$	NA	NA	NA	NA
46	InP/InGaAs HPT	1.7	NA	NA	NA	NA	NA	NA	NA
47	InP/InGaAsP HPT	1.5	12	7	$5.0\ 10^{-5}$	$1.2\ 10^{-5}$	0.5	896	0.11
48	InP/InGaAs HPT	10	12	2	$1.5\ 10^{-5}$	$1.2\ 10^{-5}$	0.1	630	0.16
This work	InGaAs/GaAsSb HPT	500	12	2	$5.0\ 10^{-10}$	$1.2\ 10^{-5}$	3.0	5.63	15.09

Table 4 Characteristics of the best reported photodetectors and this work, assuming all the detectors work at internal gain of 15 with optical power of -30 dBm and load capacitance of 2 fF.

Second, we consider the comparison to other photodetectors with internal gain (e.g. APD and HPT): GBP and other characteristics of the best reported optical receivers, needed for the energy calculation discussed in Section V, are listed in Table 4 (the first column in Table 4 corresponds to the reference numbers in the manuscript). Due to the high excess noise factor, most APDs are not able to operate at high gain, and commonly reported results in the literature are limited to a gain of ∼15. Therefore, we here use the reported currents and voltages around a gain of 15, when possible, or the maximum gain in order to be able to objectively compare a large number of the best reported devices. In addition, in this comparison we also evaluate our devices in the low-gain regime (i.e. around a gain of 15), in order to be able to compare them to the reported devices. We assumed an optical power of -30 dBm and a capacitive load of 2 fF, which represents a typical power and load (note that CMOS interconnect line capacitance can be approximated as ∼ 0.2 fF/µm, almost independent of the technology node and the typical input capacitance of most high-speed circuits is about 1 fF S13,S14). As shown in Section V, the current reaching the load is $I_{ph}/2$, and hence:

$$I_{ph}/2 = C_L \frac{dV}{dt} \tag{10}$$

Therefore, the highest current-limited bitrate for each device can be approximated by:

$$BR_{(max} \sim I_{ph}/(2C_L V_L) \tag{11}$$

The value of BR_{max}, shown in the last column in Table 4 is the lower value of this calculation (current limit) and the measured bitrate (device's bandwidth limit). Looking at the table, it is evident that the highest bitrates of integrated APDs are similar to the integrated T2-HPT when directly driving a capacitive load. When using an external amplifier, the load capacitance can be reduced and hence the bitrate increased. In the next section, we show that the bitrate of our integrated T2-HPT can be boosted to 15 Gbps with a simple amplifier, which is comparable to the bitrate of APDs with integrated amplifiers, but still with a better energy efficiency and data rate density.

7 Simulation of T2-HPT integrated with an ASIC amplifier

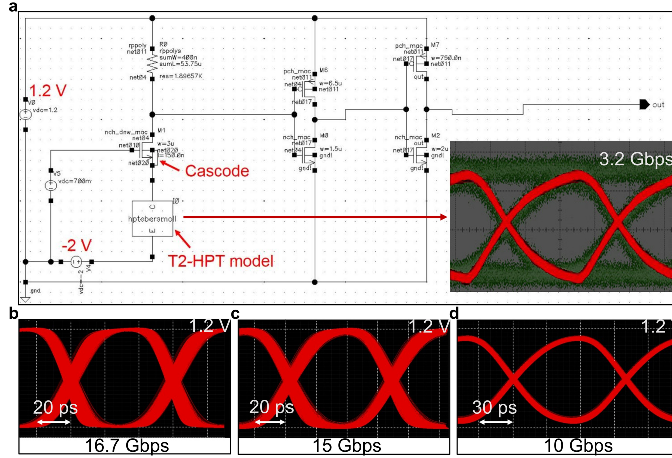

Fig. 9 Amplifier with low power consumption design for applications require high bitrate/channel. **a**, Amplifier circuit used in conjunction with the T2-HPT. The inset shows that the simulated eye diagram at 3.2 Gbps has good agreement with the measured eye diagram shown in Fig. 3a in the manuscript (the voltage swing of 20 mV matches the measured value). **b, c, d,** Simulated eye diagrams of the output of a simple amplifier integrated with the HPT detector without load (**b**), and with 5 fF (**c**) and 50 fF (**d**) load added between the output and ground, achieving open-eye at bitrates of 16.7 Gbps, 15 Gbps and 10 Gbps respectively.

The optical receiver performance of our HPT detector integrated with a 65nm ASIC was simulated using Cadence Virtuoso 6.1.8. In this circuit (Fig. 9a), a forward-active Ebers-Moll BJT model is used for the T2-HPT detector. This model shows very good agreement with our experimental results across different bias values and optical powers. A cascode transistor is then used to pin the HPT's collector voltage

near 0 V, in order to ensure proper device biasing of around 2 V. This cascode acts in conjunction with the resistor and the first inverter stage to form a simple voltage divider/ transimpedance amplifier. The second inverter stage is used to condition the signal such that this circuit can drive a larger load, in addition to adjusting the eye crossing point of the signal much closer to 50%. The CMOS portion of the circuit is powered by a standard 1.2 V supply, which shares a ground with the -2 V supply used by the detector. By adjusting the Ebers-Moll model's base-collector and base-emitter capacitance as well as the emitter resistance, eye diagrams that are almost identical to the lab measurement results at 3.2 Gbps were obtained (inset of Fig. 9a). Simulation results show that the highest data rate from the output of this circuit is 16.7 Gbps (Fig. 9b). We simulated the performance of the receiver for a moderate capacitive load of 5 fF and a large capacitive load of 50 fF, and results show data rates of 15 Gbps and 10 Gbps respectively (Fig. 9c and Fig. 9d).

References

[1] Park, M.-S., Rezaei, M., Nia, I., Brown, R., Bianconi, S., Tan, C.L., Mohseni, H.: Ingaas/inp quantum well infrared photodetector integrated on si substrate by mo/au metal-assisted wafer bonding. Optical Materials Express **8**(2), 413–419 (2018)

[2] Yu, J., Wang, Y., Lu, J.-Q., Gutmann, R.J.: Low-temperature silicon wafer bonding based on ti/ si solid-state amorphization. Applied physics letters **89**(9) (2006)

[3] Higurashi, E., Chino, D., Suga, T., Sawada, R.: Au–au surface-activated bonding and its application to optical microsensors with 3-d structure. IEEE Journal of Selected Topics in Quantum Electronics **15**(5), 1500–1505 (2009)

[4] Baraskar, A.K., Wistey, M.A., Jain, V., Singisetti, U., Burek, G., Thibeault, B.J., Lee, Y.J., Gossard, A.C., Rodwell, M.J.: Ultralow resistance, nonalloyed ohmic contacts to n-ingaas. Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena **27**(4), 2036–2039 (2009)

[5] Dayeh, S.A., Soci, C., Yu, P.K., Yu, E.T., Wang, D.: Transport properties of inas nanowire field effect transistors: The effects of surface states. Journal of Vacuum Science & Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena **25**(4), 1432–1436 (2007)

[6] Orwiler, B.: Oscilloscope vertical amplifiers. Tektronix (1969)

[7] Miller, D.A.: Physical reasons for optical interconnection. Intel J. Optoelectronics **11**, 155–168 (1997)

[8] Miller, D.A.: Device requirements for optical interconnects to silicon chips. Proceedings of the IEEE **97**(7), 1166–1185 (2009)

[9] Helme, J.P., Houston, P.A.: Analytical modeling of speed response of heterojunction bipolar phototransistors. Journal of lightwave technology **25**(5), 1247–1255 (2007)

[10] Movassaghi, Y., Fathipour, V., Fathipour, M., Mohseni, H.: Analytical and numerical evaluation of electron-injection detector optimized for swir photon detection. Journal of Applied Physics **121**(8) (2017)

[11] Bianconi, S., Rezaei, M., Park, M.-S., Huang, W., Tan, C.L., Mohseni, H.: Engineering the gain-bandwidth product of phototransistor diodes. Applied physics letters **115**(5) (2019)

[12] Fathipour, V., Mohseni, H.: Detector with internal gain for short-wave infrared ranging applications. Optical Engineering **56**(9), 091608–091608 (2017)

[13] Gasser, C., Goll, B., Zimmermann, H.: Highly sensitive hybrid optical receiver with a compact 1.4 ghz linear transimpedance amplifier in 55 nm cmos. Optical Engineering **62**(9), 095103–095103 (2023)