Article

Room temperature operation of germanium–silicon single-photon avalanche diode

https://doi.org/10.1038/s41586-024-07076-x

Received: 24 April 2023

Accepted: 16 January 2024

Published online: 21 February 2024

Check for updates

Neil Na[1]✉, Yen-Cheng Lu[1], Yu-Hsuan Liu[1], Po-Wei Chen[1], Ying-Chen Lai[1], You-Ru Lin[1], Chung-Chih Lin[1], Tim Shia[1], Chih-Hao Cheng[1] & Shu-Lu Chen[1]

The ability to detect single photons has led to the advancement of numerous research fields[1–11]. Although various types of single-photon detector have been developed[12], because of two main factors—that is, (1) the need for operating at cryogenic temperature[13,14] and (2) the incompatibility with complementary metal–oxide–semiconductor (CMOS) fabrication processes[15,16]—so far, to our knowledge, only Si-based single-photon avalanche diode (SPAD)[17,18] has gained mainstream success and has been used in consumer electronics. With the growing demand to shift the operation wavelength from near-infrared to short-wavelength infrared (SWIR) for better safety and performance[19–21], an alternative solution is required because Si has negligible optical absorption for wavelengths beyond 1 μm. Here we report a CMOS-compatible, high-performing germanium–silicon SPAD operated at room temperature, featuring a noise-equivalent power improvement over the previous Ge-based SPADs[22–28] by 2–3.5 orders of magnitude. Key parameters such as dark count rate, single-photon detection probability at 1,310 nm, timing jitter, after-pulsing characteristic time and after-pulsing probability are, respectively, measured as 19 kHz μm^{-2}, 12%, 188 ps, ~90 ns and <1%, with a low breakdown voltage of 10.26 V and a small excess bias of 0.75 V. Three-dimensional point-cloud images are captured with direct time-of-flight technique as proof of concept. This work paves the way towards using single-photon-sensitive SWIR sensors, imagers and photonic integrated circuits in everyday life.

Detectors that are extremely sensitive and may discern the smallest quanta of light—that is, a single photon—are referred to as single-photon detectors. These detectors have played crucial parts during the past decades in scientific research on optical quantum computation[1,2], quantum key distribution[3,4], quantum imaging[5,6] and biophotonics[7], by detecting the granular properties of light, and in recent technological developments of light detection and ranging (LIDAR) in a coaxial setting with single detectors[8,9], or in-camera setting with pixel arrays[10,11], by time-resolving the photon arrival events. Solid-state implementations of these detectors through superconductors, for example, superconducting nanowire single-photon detector (SNSPD), and semiconductors, for example, single-photon avalanche diode (SPAD) or sometimes referred to as Geiger-mode avalanche photodiode (APD), are perhaps the most well-known examples. Although SNSPD detects a broad wavelength range from mid-infrared to visible, it typically operates at a cryogenic temperature of a few K (refs. 12–14), which limits its use mainly as a lab instrument. By contrast, SPAD based on InGaAs or Si detects only a specific wavelength range from SWIR to near infrared (NIR) or from NIR to visible, respectively, but it can be operated at near room temperature or room temperature[12,15–18] and hence has become popular for general use. In particular, Si-based SPAD is intrinsically CMOS compatible, in which homogenous integration with CMOS circuits (that is, monolithic) at sub-micron technology node under front-side illumination[17] as well as heterogenous integration with CMOS circuits (that is, wafer-to-wafer bonding) at deep sub-micron technology node under backside illumination[18] have been demonstrated. These advancements have enabled Si-based SPAD to enter the segment of mainstream consumer electronics as proximity sensors and three-dimensional (3D) imagers at NIR wavelength. Recently, SWIR has emerged as the next-generation choice of wavelength because of the following reasons: (1) better laser eye safety as NIR can cause damage to the retina[19]; (2) lower ambient light interference as solar spectrum is weaker at SWIR or even absent at around 1,380 nm wavelength[20]; (3) higher atmospheric transmission as free-space optical transmissivity is dominated by Rayleigh scattering[21]; and (4) less interference with Si devices and circuits as Si absorbs NIR but not SWIR. Consequently, SWIR single-photon detectors operated at room temperature and being CMOS compatible are actively sought after. In the literature, Ge-based SPAD is shown as a potential candidate, but it experimentally shows a high dark count rate (DCR) and/or a low single-photon detection probability (SPDP) so that notable performance can be reached only at a low temperature typically less than 200 K (refs. 22–28), or it is studied only through theoretical simulations without experimental supports[29,30].

Benchmark and concept of the GeSi SPAD device

In this study, we report the fabrication of a 15-μm diameter germanium–silicon (GeSi) SPAD with a CMOS-compatible fabrication process

[1]Artilux Inc., Zhubei, Taiwan ROC. ✉e-mail: neil@artiluxtech.com

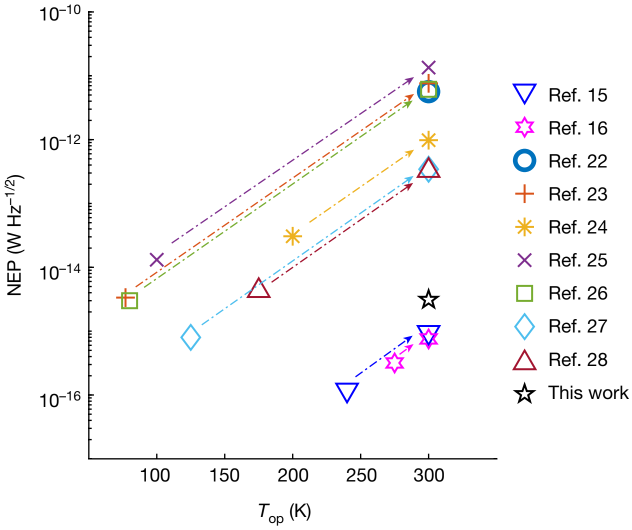


Fig. 1 | NEP and operation temperature of the previously reported Ge-based[22–28] and InGaAs-based[15,16] SPADs. The dash-dotted arrows indicate the NEP of the previously reported SPADs if they were operated at 300 K, assuming DCR is doubled for every 10 K increase in temperature. Note that all devices are tested at 1,310 nm wavelength except for ref. 22, which is tested at 1,210 nm. The device diameters are 25 µm (ref. 15), 10 µm (ref. 16), 120 µm (ref. 22), 40 µm (ref. 23), 30 µm (ref. 24), 25 µm (ref. 25), 4.5 µm (effective) (ref. 26), 100 µm (ref. 27), 26 µm (ref. 28) and 15 µm (this work).

and characterize it at room temperature to demonstrate its high performance. To benchmark the demonstrated GeSi SPAD with the Ge-based and InGaAs-based SPADs in the literature, we calculate the noise-equivalent power (NEP) by using the definition[31]

$$\mathrm{NEP} = \frac{\hbar\omega\sqrt{2\times \mathrm{DCR}}}{\mathrm{SPDP}}, \quad (1)$$

where the DCR and SPDP from the highest operation voltages at the highest operation temperatures are used in the calculations. $\hbar$ is the reduced Plack's constant and ω is the angular frequency of the photon. In Fig. 1, the operation temperature and NEP in refs. 22–28 (Ge-based), refs. 15,16 (InGaAs-based) and this work are plotted. To better compare the NEP spread over wide ranges of temperatures, without loss of generality, DCR is assumed to double for every 10 K increase in temperature[31], which is typical for Ge-based and InGaAs-based devices. Then, the dash-dotted arrows in Fig. 1 are added to indicate the NEP of the previously reported Ge-based and InGaAs-based SPADs if they were operated at 300 K: the demonstrated GeSi SPAD features a low NEP around a few fW Hz$^{-1/2}$ at room temperature, which is about 2–3.5 orders of magnitude lower than those of the previously reported Ge-based SPADs. The marked difference is mainly because of the low dark current of the demonstrated GeSi SPAD. Taking ref. 24 as an example, which is one of the first demonstration of Ge-on-Si separate-absorption-charge-multiplication APD for SPAD, it features a dark current of about 1 µA at the punch-through point around −22 V, whereas the demonstrated GeSi SPAD features a dark current of about 20 pA at the punch-through point around −9 V. Note that in Fig. 1 the device areas are not normalized, because for the previously reported Ge-based SPADs the DCR depends largely on design and process and does not scale well with the device area (for example, in ref. 27 the area is about 15 times larger than that in ref. 28 but their NEP values are about the same). Still the demonstrated GeSi SPAD has an NEP that is on average 2.5 orders of magnitude lower than those of the previously reported Ge-based SPADs if the device areas are normalized. By contrast, the NEP of the demonstrated GeSi SPAD is close to the NEP of the previously reported InGaAs-based SPADs with a difference of about only 0.5 orders of magnitude. Moreover, the after-pulsing characteristic time of the demonstrated GeSi SPAD is much shorter than those of the previously reported InGaAs-based SPADs, suggesting that for certain applications the demonstrated GeSi SPAD may already outperform some of the previously reported InGaAs-based SPADs (elaborated later in the text).

The notable reduction in the dark current of the demonstrated GeSi SPAD is made possible by addressing the following issues: (1) the threading dislocations occurring on the (311) planes in Ge arising from the 4% lattice mismatch between Ge and Si; (2) the lack of high-quality surface passivation agents over Ge surfaces; and (3) the high electric fields present in the low-quality material regions, especially along the Ge sidewalls. For the issue of threading dislocation, various methods have been followed such as graded buffer layers grown between Si and Ge (ref. 32), post-Ge high-temperature thermal annealing[33], low-energy plasma-enhanced chemical vapour deposition[34], high-aspect-ratio trapping of defects[35] and Ge-on-insulator with defects chemically and mechanically polished[36], to name a few. For the issue of surface passivation, various agents have been experimented such as amorphous Si (ref. 37), GeO_2 (ref. 38), Al_2O_3 (ref. 36) and pure Ga/B (ref. 39), to name a few. Generally, with the combined efforts of reducing threading dislocations and surface dangling bonds, Ge-on-Si photodiode featuring a low dark current density down to hundreds or even tens of µA cm^{-2} can be accomplished at a low reverse-bias voltage. However, to be used as a high-performing APD or even SPAD, such a low dark current must be maintained while the device is biased above the unity gain point, at which the high electric fields needed for transporting photocarriers and generating impact ionizations begin to take effect. The previously reported Ge-based APDs[40–45] did not meet this criterion, and here the demonstrated GeSi SPAD achieves this criterion through a previously unknown design to effectively screen the high electric fields off the low-quality Ge sidewalls. Figure 2a shows the schematic concept of the demonstrated GeSi SPAD. Figure 2b shows the corresponding scanning electron microscope (SEM) image. The device consists of elements such as the p-contact doping region (deepest blue) in Ge, the p-type charge layer (lighter blue) in Si and the deep n-well doping region connected to the n-contact doping region through the n-sinker doping region all in Si. Then, a p-type screening layer (deeper blue) in Si is implemented at regions above the charge layer and below the Ge sidewalls (for more fabrication details, see the Methods). The dopant density of the screening layer must be high enough so that when the charge layer is depleted, the high electric fields below the screening layer would terminate at the screening layer, that is, would not penetrate the Ge sidewalls. At the same time, the dopant density of the screening layer must be low enough so that premature breakdown below the screening layer would not occur, that is, the breakdown electric fields below the charge layer are uniform regardless of whether the screening layer above presents or not. Through electrostatic simulation, we observe negligible electrical fields much less than 1 kV cm^{-1} along the Ge sidewalls even when the device is biased far above its breakdown voltage.

GeSi SPAD device characterization

Figure 2c shows the I–V measurement from wafer-level testing in which the reverse-bias current I is plotted as a function of the applied voltage V_{ap}, with and without a collimated 1,310 nm laser beam to illuminate the device under testing (DUT). Two DUT samples S1 and S2 of 15 µm diameter each are used, and the charge layer concentration of S1 is slightly larger than that of S2. A reference photodiode, that is, the same 15-µm diameter device in which the charge layer is removed, shows a photocurrent of about 1 µA under the same measurement conditions. By using this number and intercepting the photocurrents in Fig. 2c, the corresponding primary (that is, at unity gain point) dark currents of S1 and S2 are determined to be about the same 8 pA at −9.2 V and

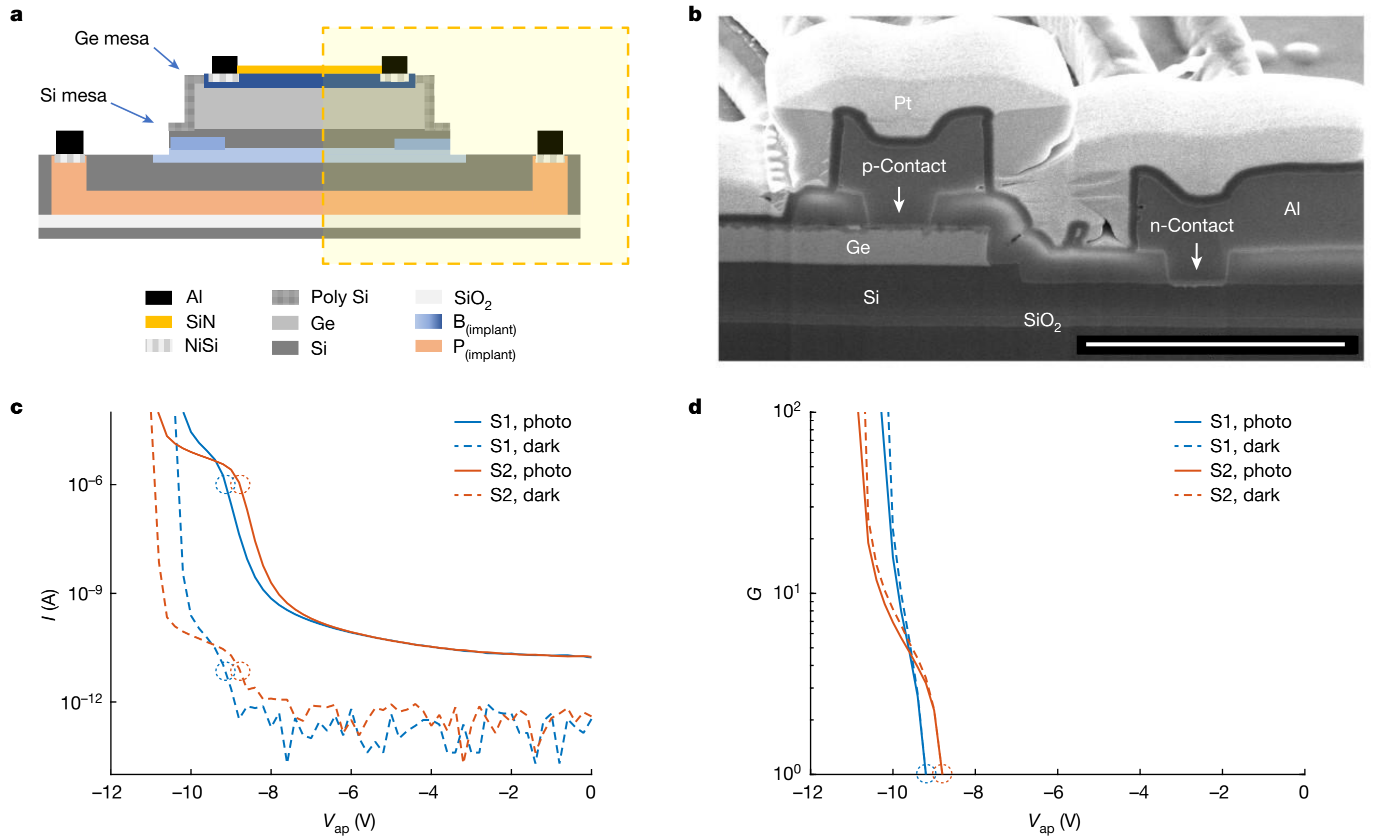


Fig. 2 | Cross-sectional schematic concept, SEM image and *I*–*V* characteristics of the demonstrated GeSi SPAD. a, Schematic concept of the demonstrated GeSi SPAD. Boron implants are used to create the p-contact region (deepest blue) in Ge, and phosphorus implants are used to create the n-contact region in Si. Additional boron implants in Si form two regions below Ge, one as the charge layer (lighter blue) and the other as the screening layer (deeper blue). Additional phosphorus implants in Si form the deep n-well region above the buried SiO_2, and the n-sinker region between the n-contact region and the deep n-well region. **b**, SEM image. The image is captured corresponding to the yellow area enclosed by the yellow dashed lines in **a**. Note that Pt, the white conformal layer, is deposited over the device to avoid SEM charging. **c**,**d**, Photocurrents (solid curves) and dark currents (dashed curves) as well as their corresponding gains, plotted as a function of applied voltage, for samples S1 (blue) and S2 (red). By intercepting the photocurrents with the photocurrent of a reference photodiode, the primary photocurrents and dark currents in **c** and the unity gain points in **d** for samples S1 and S2 are, respectively, identified and labelled by the blue and red dotted circles. Scale bar, 3 µm (**b**).

−8.8 V, respectively, corresponding to 4.5 µA cm^{-2} and is the lowest value reported, to our knowledege, for Ge-based APD and/or SPAD at room temperature. Additional *I*–*V* measurements show that the primary dark current is increased by about 12 times when the device diameter is increased from 15 µm to 50 µm, suggesting that the primary dark current scales reasonably well with the device area. The punch-through voltage V_{PT} of S1 and S2 are determined to be −9.3 V and −9 V, respectively, from additional *C*–*V* measurements and coinciding with the 'knee' points of the photocurrents in Fig. 2c. The breakdown voltage V_{BD} of S1 and S2 are determined to be −10.26 V and −10.86 V, respectively, by defining 10 µA dark current as the breakdown point. Moreover, the breakdown voltage thermal coefficient, defined as $(\Delta V_{BD}/V_{BD})/\Delta T$, where ΔV_{BD} and ΔT are the increments of breakdown voltage and temperature, is determined from additional temperature-dependent *I*–*V* measurements to be about 0.06% per °C between 30 °C and 80 °C, which agrees reasonably with the case of Si PIN APD[46]. The voltage differences between V_{BD} and V_{PT} for S1 and S2 are 1 V and 1.9 V, respectively, which show the difference in charge layer concentrations. It can be observed that the shapes of the photo *I*–*V* curve and the corresponding dark *I*–*V* curve in Fig. 2c share marked similarities, so we plot the photocurrent gain and the corresponding dark current gain as a function of the applied voltage in Fig. 2d for comparison. It is found that they overlap with each other up to a gain of about 100, suggesting that the multiplied component of the dark current originates at places away from the low-quality Ge sidewalls because of the screening layer.

To characterize the performance of the demonstrated GeSi SPAD, gated-mode operation[22–28] is used in which the cathode of the DUT is biased repeatedly at repetition time T_r through a bias-Tee with a direct-current (d.c.) voltage below V_{BD} and a radiofrequency (RF) voltage pulse causing a total operation voltage above V_{BD} during gate time T_g (for more measurement details, see the Methods). The count rate directly reported by the counter or the oscilloscope should be multiplied by T_r/T_g to obtain the actual DCR. As T_g is typically around a few or tens of ns, DCR below GHz or sub-GHz range can be time-resolved and hence gated-mode operation has been used in characterizing SPADs that may have wide ranges of DCRs. In Fig. 3a, the measured DCR from die-level testing using the counter is plotted as a function of excess bias voltage V_{ex} (the difference voltage between the operation voltage above V_{BD} and V_{BD}) for S1 and S2, given T_g and T_r equal to 5 ns and 102.4 ns, respectively. The DCR between 1.5 MHz and 14.1 MHz (about 8.5 kHz µm^{-2} and 79.8 kHz µm^{-2}) can be reached for V_{ex} between 0.5 V and 2.5 V (about 2% and 23% of V_{BD}), which are the lowest values reported, to our knowledge, for Ge-based SPAD at room temperature. Next, 1,310 nm picosecond laser pulses in synchronization with gate pulses are injected into a confocal microscope setup that couples to the device (for more measurement details, see the Methods). The average number of photons per pulse ⟨*n*⟩ received by the DUT is controlled by first calibrating the optical loss of coupling the confocal microscope to the device and then adjusting a variable optical attenuator. SPDP can be obtained by subtracting the measured DCR from the measured total

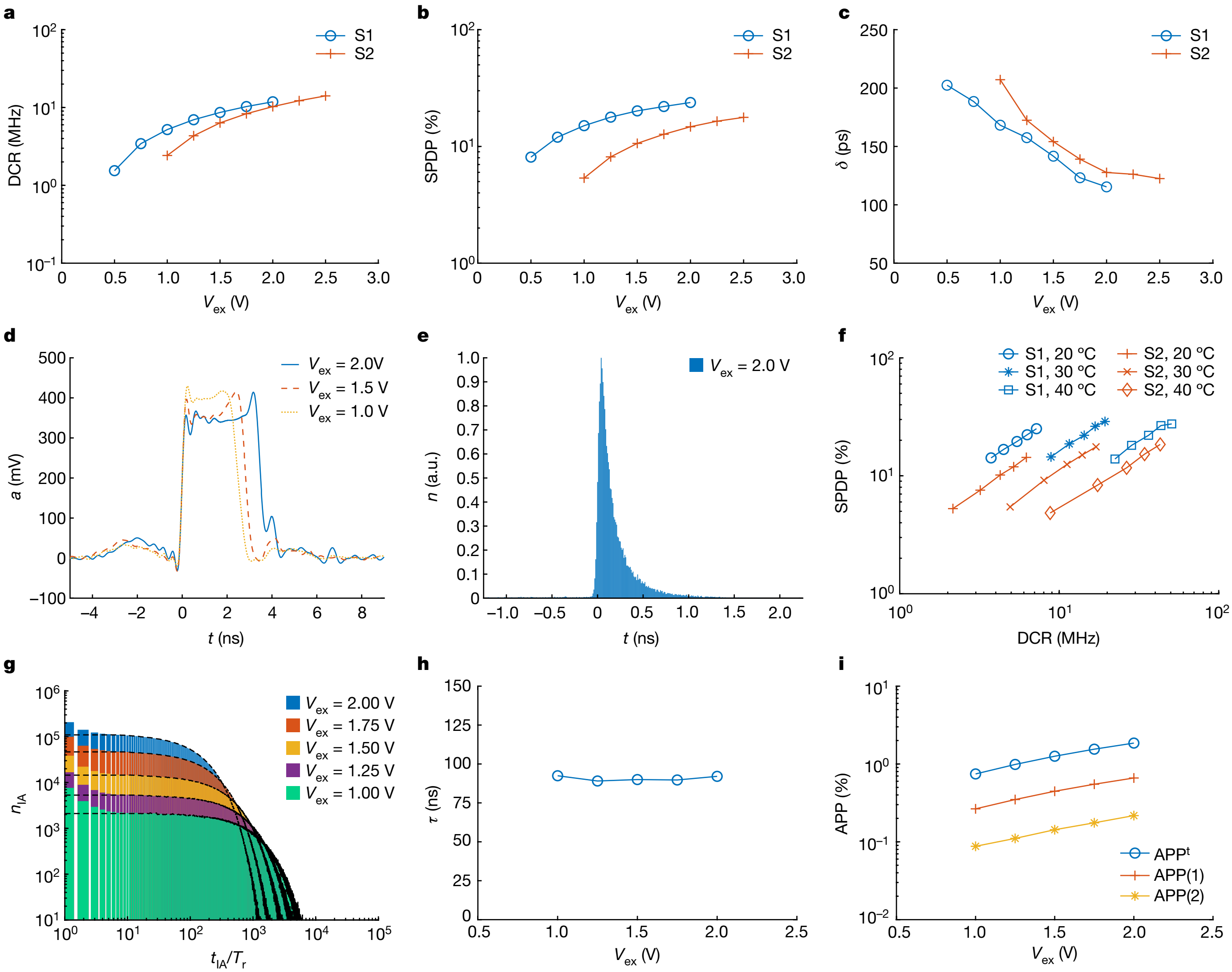


Fig. 3 | Counting properties of the demonstrated GeSi SPAD. a, DCR measured as a function of V_{ex} for samples S1 and S2. **b**, SPDP measured as a function of V_{ex} for samples S1 and S2, given ⟨n⟩ = 0.1. **c**, Timing jitter δ in FWHM measured as a function of V_{ex} for samples S1 and S2, given ⟨n⟩ = 10. **d**, Exemplary voltage waveforms from the DCR measurement of sample S1 plotted as a function of time. **e**, Exemplary photon arrival histogram from the δ measurement of sample S1 plotted as a function of time. **f**, Temperature-dependent DCR and SPDP measured as a function of V_{ex} equal to 1 V, 1.25 V, 1.5 V, 1,75 V and 2 V for samples S1 and S2, for chuck temperatures of 20 °C, 30 °C and 40 °C. **g**, Interarrival count n_{IA} plotted as a function of interarrival time t_{IA} normalized by T_r. **h**, After-pulsing characteristic time τ calculated with the equations in the Methods, and plotted as a function of V_{ex}. **i**, APP calculated with the equations in the Methods, and plotted as a function of V_{ex}. In all cases, T_g and T_r are set to be 5 ns and 102.4 ns, respectively.

count rate (TCR), that is, the count rate with the pulsed laser turned on, and multiplying it by $T_g/\langle n\rangle$ assuming $\langle n\rangle \ll 1$ to ensure the probability of a laser pulse containing multiple photons is negligible. In Fig. 3b, the measured SPDP from die-level testing using the counter is plotted as a function V_{ex} for S1 and S2, given the same conditions of T_g and T_r as in Fig. 3a. SPDP between 5% and 24% can be reached for V_{ex} between 0.5 V and 2.5 V given ⟨n⟩ = 0.1. In Fig. 3c, the measured timing jitter δ in full-width-half-maximum (FWHM) from die-level testing using the oscilloscope is plotted as a function V_{ex} for S1 and S2, given the same conditions of T_g and T_r as in Fig. 3b. Timing jitter δ between 115 ps and 207 ps can be reached for V_{ex} between 0.5 V and 2.5 V given ⟨n⟩ = 10. Note the condition ⟨n⟩ = 10 is chosen for satisfying a signal-to-background-noise ratio equal to 10 (assuming SPDP ≈ 10%, DCR ≈ 2 MHz and T_g ≈ 5 ns) commonly required for single-shot LIDAR. An exemplary voltage waveforms from the DCR measurement of S1 in Fig. 3a operated at V_{ex} equal to 1 V, 1.5 V and 2 V are shown in Fig. 3d, in which sharp signal rise due to breakdown and sharp signal fall due to turning gate on and off are observed. An exemplary photon arrival histogram from the δ measurement of S1 in Fig. 3c operated at V_{ex} equal to 2 V is shown in Fig. 3e, in which δ of 115 ps is observed and attributed mainly to the device δ because the oscilloscope system δ is estimated to be around a few tens of picoseconds and is negligible. Note the shape of the histogram is exponentially decreasing with time because the device δ is dominated by the long RC time constant of the device (roughly between 100 ps and 200 ps) because of higher contact resistance and higher parasitic capacitance. Finally, as shown in Fig. 3f, DCR and SPDP of S1 and S2 are measured from wafer-level testing using the counter at V_{ex} equal to 1 V, 1.25 V, 1.5 V, 1.75 and 2 V, when the chuck is controlled to 20 °C, 30 °C and 40 °C. For DCR, its increase with increasing temperature is about 2.5 times for every 10 °C, corresponding to an activation energy of about 0.64 eV and is slightly smaller than the Ge bandgap of about 0.66 eV. This suggests that defects in Ge with energy levels close to Ge band edges are responsible for generating DCR, and so DCR can be further minimized by optimizing the fabrication process. As for SPDP, it stays roughly the same throughout the controlled temperature range.

When a breakdown event occurs and induces strong current flow passing through the multiplication region, some carriers may be trapped in the defects and subsequently released at a time delay causing another breakdown event. This phenomenon is called after-pulsing, and its characteristic time τ and occurrence after-pulsing probability (APP) are crucial parameters in SPAD operations. Here we characterize the after-pulsing of the demonstrated GeSi SPAD at room temperature, by applying time-correlated counting under gated-mode operation to measure the statistics of interarrival time with dark counts[47,48] (for more measurement and calculation details, see the Methods). In Fig. 3g, the measured histograms for interarrival time normalized by T_r are shown for several V_{ex}, given the same conditions of T_g and T_r as in Fig. 3a–f. The black dashed lines are fitting curves for the thermally generated counts $n_{TG}(i)$, whereas the after-pulsing counts $n_{AP}(i)$ are above the fitting curves and distribute exponentially, where i is the index representing the ith gate time after the reference gate time. In Fig. 3h, the calculated τ is shown as a function of V_{ex}. It is observed that τ is relatively insensitive to V_{ex} and is about 90 ns for V_{ex} between 1 V and 2 V. In Fig. 3i, the calculated APP(1) (that is, $i = 1$), APP(2) (that is, $i = 2$), and APP^t (that is, the total APP) are shown as a function of V_{ex}. It is observed that APP^t is exponentially related to V_{ex} and is between 0.7% and 1.9% for V_{ex} between 1 V and 2 V. The numbers of τ and APP are found to be in reasonable agreement with those of the Si-based SPADs at room temperature[49], suggesting the high crystalline quality of the demonstrated GeSi SPAD. Note for certain applications such as multi-shot LIDAR based on direct time-of-flight (TOF) technique using pulse rather than continuous wave, to reach a target signal-to-background-noise ratio, the square root of the integration time required is proportional to NEP$\sqrt{\tau}$. Consequently, the performance of the demonstrated GeSi SPAD may already outperform some of the previously reported InGaAs-based SPADs (considering NEP about 3.5 times lower from Fig. 1; $\sqrt{\tau}$ about 3–33 times higher from refs. 15,16) in this and other similar cases.

LIDAR system demonstration

As proof of concept, the setup for producing data shown in Fig. 3 is modified to apply the demonstrated GeSi SPAD for LIDAR based on direct TOF technique. 3D scanning and imaging are done by replacing the variable optical attenuator with a coaxial transceiver that uses x–y galvo-mirrors to steer the fired laser pulses and collect the returned laser pulses along the same optical path (for more measurement details, see the Methods). As shown by the two-dimensional RGB (red, green and blue) image in Fig. 4a, the 3D objects are placed against the wall at different distances away from the mirrors under normal indoor ceiling lighting. Alphabet letters standing on the lower cardboard staircase are 100 mm in height, and the platonic solids standing on the upper cardboard staircase are 65 mm in height. The captured 3D point cloud (PCL) image by using sample S1 at V_{ex} equal to 2 V is shown in Fig. 4b. The relative distances away from the mirrors are colour-coded for better visualization, and it can be seen that the spatial orientations of the 3D objects are nicely revolved without any post-processing, attesting to the potential of the demonstrated GeSi SPAD for LIDAR based on direct TOF technique. Two TOF histograms recorded for paths P1 and P2 are shown in Fig. 4c, corresponding to the paths that scatter the laser pulses at locations indicated by the two blue circles and the one red circle marked in Fig. 4b. $\langle n \rangle$ received by the GeSi SPAD are estimated to be 0.02 and 1.5 for P1 and P2, respectively. δ values of the two peaks for P1 and the one peak for P2, after removing the background dark counts and counter system δ around 99 ps, are found to be about the same 294 ps, which is larger than the δ shown in Fig. 3e with $\langle n \rangle = 10$ by a factor of 2.56.

Summary and outlook

We have demonstrated the Geiger-mode operation of a high-performing GeSi APD as SPAD at room temperature, which was previously deemed not possible and shall be a strong contender for room-temperature-operated, CMOS-compatible, SWIR single-photon detectors. As Ge-based detector operated in linear mode has been commercialized through Si photonic platform[37] and CMOS image sensor plaform[50], the ability of operating in Geiger-mode shall boost the developments of optoelectronic devices and circuits over these platforms towards unprecedented single-photon level sensitivity.

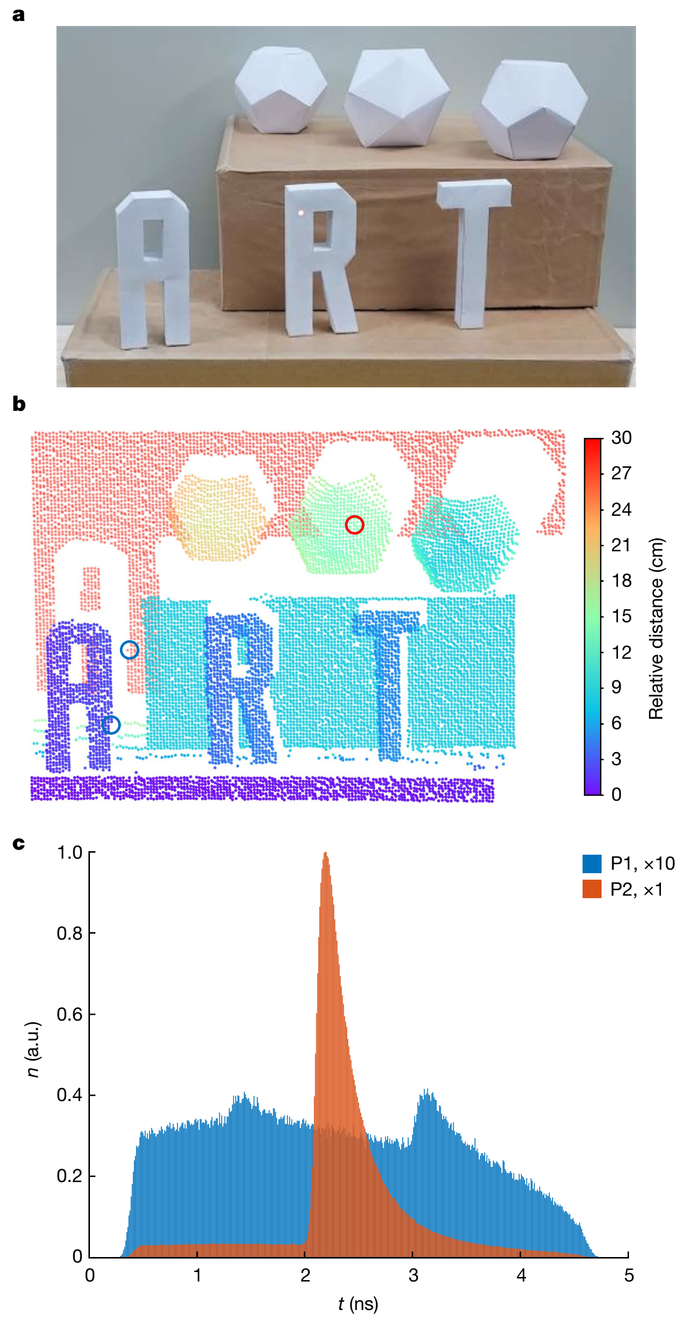


Fig. 4 | Experimental demonstration of 3D scanning and imaging. a, The RGB image of the scene. The red laser spot on the alphabet letter R marks the initial point of steering galvo-mirrors. **b**, The PCL image of the scene, captured by the demonstrated GeSi SPAD with direct TOF technique. The colour bar indicates the relative distance away from the mirrors in cm. **c**, The normalized TOF histograms for paths P1 and P2, in which P1 and P2 correspond to the paths that scatter the 1,310 nm laser pulses at locations marked in **b** with two blue circles and one red circle, respectively. The normalized TOF histogram for P1 is scaled 10 times so that P1 and P2 can be visually compared more easily. a.u., arbitrary units.

Future directions include (1) scaling the device area for developing a sensing and imaging pixel with a pitch smaller than 10 μm; (2) minimizing the contact resistance and the parasitic capacitance; and (3) heterogenous integration with CMOS circuits through wafer-level bonding.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-024-07076-x.

1. Politi, A., Matthews, J. C. F. & O'Brien, J. L. Shor's quantum factoring algorithm on a photonic chip. *Science* **325**, 1221 (2009).
2. Zhong, H.-S. et al. Quantum computational advantage using photons. *Science* **370**, 1460–1463 (2020).
3. Waks, E. et al. Quantum cryptography with a photon turnstile. *Nature* **420**, 762 (2002).
4. Lai, S.-K. et al. Satellite-relayed intercontinental quantum network. *Phys. Rev. Lett.* **120**, 030501 (2018).
5. Brida, G., Genovese, M. & Ruo Berchera, I. Experimental realization of sub-shot-noise quantum imaging. *Nat. Photon.* **4**, 227–230 (2010).
6. Zarghami, M. et al. A 32×32-pixel CMOS imager for quantum optics with per-SPAD TDC, 19.48% fill-factor in a 44.64-μm pitch reaching 1-MHz observation rate. *IEEE J. Solid-State Circuits* **55**, 2819–2830 (2020).
7. Bruschini, C., Homulle, H., Antolovic, I. M., Burri, S. & Charbon, E. Single-photon avalanche diode imagers in biophotonics: review and outlook. *Light Sci. Appl.* **8**, 87 (2019).
8. Pawlikowska, A. M., Halimi, A., Lamb, R. A. & Buller, G. S. Single-photon three-dimensional imaging at up to 10 kilometers range. *Opt. Express* **25**, 11919–11931 (2017).
9. Kuzmenko, K. et al. 3D LIDAR imaging using Ge-on-Si single-photon avalanche diode detectors. *Opt. Express* **28**, 1330–1344 (2020).
10. Entwistle, M. et al. Geiger-mode APD camera system for single-photon 3D LADAR imaging. In *Proc. SPIE* Vol. 8375, Advanced Photon Counting Techniques VI, 83750D (SPIE, 2012).
11. Morimoto, K. et al. Megapixel time-gated SPAD image sensor for 2D and 3D imaging applications. *Optica* **7**, 346–354 (2020).
12. Hadfield, R. H. Single-photon detectors for optical quantum information applications. *Nat. Photon.* **3**, 696–705 (2009).
13. Marsili, F. et al. Detecting single infrared photons with 93% system efficiency. *Nat. Photon.* **7**, 210–214 (2013).
14. Najafi, F. et al. On-chip detection of non-classical light by scalable integration of single-photon detectors. *Nat. Commun.* **6**, 5873 (2015).
15. Campbell, J. C., Sun, W., Lu, Z., Itzler, M. A. & Jiang, X. Common-mode cancellation in sinusoidal gating with balanced InGaAs/InP single photon avalanche diodes. *IEEE J. Quantum Electron.* **47**, 1505–1511 (2012).
16. Signorelli, F. et al. Low-noise InGaAs/InP single-photon avalanche diodes for fiber-based and free-space applications. *IEEE J. Sel. Top. Quantum Electron.* **28**, 3801310 (2022).
17. Xu, H., Pancheri, L., Dalla Betta, G.-F. & Stoppa, D. Design and characterization of a p+/n-well SPAD array in 150nm CMOS process. *Opt. Express* **25**, 12765–12778 (2017).
18. Lee, M.-J. et al. High-performance back-illuminated three-dimensional stacked single-photon avalanche diode implemented in 45-nm CMOS technology. *IEEE J. Sel. Top. Quant. Electron.* **24**, 3801809 (2018).
19. Sliney, D. H. & Freasier, B. C. Evaluation of optical radiation hazards. *Appl. Opt.* **12**, 1–24 (1973).
20. Bird, R. E., Hulstrom, R. L. & Lewis, L. J. Terrestrial solar spectral data sets. *Sol. Energy* **30**, 563–574 (1983).
21. Arnulf, A., Bricardi, J., Cura, E. & Varet, C. Transmission by haze and fog in the spectral region 0.35 to 10 microns. *J. Opt. Soc. Am.* **47**, 491–498 (1957).
22. Loudon, A. Y. et al. Enhancement of the infrared detection efficiency of silicon photon-counting avalanche photodiodes by use of silicon germanium absorbing layers. *Opt. Lett.* **27**, 219–221 (2002).
23. Tosi, A., Dalla Mora, A., Zappa, F. & Cova, S. Germanium and InGaAs/InP SPADs for single-Photon detection in the near-infrared. In *Proc. SPIE* Vol. 6771, Advanced Photon Counting Techniques II, 67710P (SPIE, 2007).
24. Lu, Z. et al. Geiger-mode operation of Ge-on-Si avalanche photodiodes. *IEEE J. Quantum Electron.* **47**, 731–735 (2011).
25. Warburton, R. E. et al. Ge-on-Si single-photon avalanche diode detectors: design, modeling, fabrication, and characterization at wavelengths 1310 and 1550 nm. *IEEE Trans. Electron Devices* **60**, 3807–3813 (2013).
26. Martinez, N. J. D. et al. Single photon detection in a waveguide-coupled Ge-on-Si lateral avalanche photodiode. *Opt. Express* **25**, 161310–16139 (2017).
27. Vines, P. et al. High performance planar germanium-on-silicon single-photon avalanche diode detectors. *Nat. Commun.* **10**, 1086 (2019).
28. Ferre Llin, L. et al. High sensitivity Ge-on-Si single-photon avalanche diode detectors. *Opt. Lett.* **45**, 6406–6409 (2020).
29. Ke, S. et al. Design of wafer-bonded structures for near room temperature Geiger-mode operation of germanium on silicon single-photon avalanche photodiode. *Appl. Opt.* **56**, 4646–4653 (2017).
30. Soref, R. A., De Leonardis, F. & Passaro, V. M. N. Simulations of nanoscale room temperature waveguide-coupled single-photon avalanche detectors for silicon photonic sensing and quantum applications. *ACS Appl. Nano Mater.* **2**, 7503–7512 (2019).
31. Thorburn, F. et al. Ge-on-Si single-photon avalanche diode detectors for short-wave infrared wavelengths. *J. Phys. Photonics* **4**, 012001 (2022).
32. Samavedam, S. B., Currie, M. T., Langdo, T. A. & Fitzgerald, E. A. High-quality germanium photodiodes integrated on silicon substrates using optimized relaxed graded buffers. *Appl. Phys. Lett.* **73**, 2125–2127 (1998).
33. Luan, H.-C. et al. High-quality Ge epilayers on Si with low threading-dislocation densities. *Appl. Phys. Lett.* **75**, 2909–2911 (1999).
34. Osmond, J. et al. Ultralow dark current Ge/Si(100) photodiodes with low thermal budget. *Appl. Phys. Lett.* **94**, 201106 (2009).
35. Åberg, I. et al. A low dark current and high quantum efficiency monolithic germanium-on-silicon CMOS imager technology for day and night imaging applications. In *Proc. 2010 International Electron Devices Meeting (IEDM)* 14.4.1–14.4.4 (IEEE, 2010).
36. Lee, K.-H. et al. Defects reduction of Ge epitaxial film in a germanium-on-insulator wafer by annealing in oxygen ambient. *APL Mater.* **3**, 016102 (2015).
37. Morse, M., Dosunmu, O., Sarid, G. & Chetrit, Y. Performance of Ge-on-Si p-i-n photodetectors for standard receiver modules. *IEEE Photon. Technol. Lett.* **18**, 2442–2444 (2006).
38. Takenaka, M., Morii, K., Sugiyama, M., Nakano, Y. & Takagi, S. Dark current reduction of Ge photodetector by GeO_2 surface passivation and gas-phase doping. *Opt. Express* **20**, 8718–8725 (2012).
39. Sammak, A., Aminian, M., Nanver, L. K. & Charbon, E. CMOS-compatible PureGaB Ge-on-Si APD pixel arrays. *IEEE Trans. Electron Devices* **63**, 92–99 (2016).
40. Kang, Y. et al. Monolithic germanium/silicon avalanche photodiodes with 340 GHz gain–bandwidth product. *Nat. Photon.* **3**, 59–63 (2008).
41. Duan, N., Liow, T.-Y., Lim, A. E.-J., Ding, L. & Lo, G. Q. 310 GHz gain-bandwidth product Ge/Si avalanche photodetector for 1550 nm light detection. *Opt. Express* **20**, 11031–11036 (2012).
42. Huang, M. et al. Germanium on silicon avalanche photodiode. *IEEE J. Sel. Top. Quantum Electron.* **24**, 3800911 (2018).
43. Assefa, S., Xia, F. & Vlasov, Y. A. Reinventing germanium avalanche photodetector for nanophotonic on-chip optical interconnects. *Nature* **464**, 80–84 (2010).
44. Virot, L. et al. Germanium avalanche receiver for low power interconnects. *Nat. Commun.* **5**, 4957 (2014).
45. Chen, H. T. et al. High sensitivity 10Gb/s Si photonic receiver based on a low-voltage waveguide-coupled Ge avalanche photodetector. *Opt. Express* **23**, 815–822 (2015).
46. Su, Y. K., Chang, C. Y. & Wu, T. S. Temperature dependent characteristics of a PIN avalanche photodiode (APD) in Ge, Si and GaAs. *Opt. Quantum Electron.* **11**, 109–117 (1979).
47. Hofbauer, M., Steind, B. & Zimmermann, H. Temperature dependence of dark count rate and after pulsing of a single-photon avalanche diode with an integrated active quenching circuit in 0.35 μm CMOS. *J. Sens.* **2018**, 9585931 (2018).
48. Mahmoudi, H., Hofbauer, M., Goll, B. & Zimmermann, H. Noise and breakdown characterization of SPAD detectors with time-gated photon-counting operation. *Sensors* **21**, 5287 (2021).
49. Ziarkash, A. W., Joshi, S. K., Stipčević, M. & Ursin, R. Comparative study of afterpulsing behavior and models in single photon counting avalanche photo diode detectors. *Sci. Rep.* **8**, 5076 (2018).
50. Na, N. et al. High-performance germanium-on-silicon lock-in pixels for indirect time-of-flight applications. In *Proc. 2018 International Electron Devices Meeting (IEDM)* 32.4.1–32.4.4 (2018).

Methods

Fabrication

The device fabrication starts with phosphorus implants over a silicon-on-insulator (SOI) wafer to form the deep n-well region and the lower part of the n-sinker region in Si, followed by an epitaxial growth of 250 nm Si (multiplication layer) on the SOI wafer with ultrahigh-vacuum chemical vapour deposition. The buried SiO_2 of the SOI wafer is used to isolate the leakage current from the bulk Si substrate. Boron implants form the charge layer and the screening layer around 2.5 µm in width in Si, and phosphorus implants form the upper part of the n-sinker and the n-contact regions in Si, followed by another epitaxial growth of buffer Si and 450 nm Ge (absorption layer) between 300 °C and 600 °C. After a post-epitaxy thermal annealing between 600 °C and 800 °C, a series of dry and wet etches form a Ge mesa and a Si mesa, in which polycrystalline-Si-based passivation is processed over the Ge mesa. Boron implants form the p-contact region in Ge, followed by the Ni silicide process to create the ohmic contacts over both n-contact and p-contact regions, and the Al metallization process to define the electrical traces and pads for electrical probing. Finally, SiN is deposited and etched to form an anti-reflection coating layer on top of the Ge mesa for optical illumination.

Measurement

The electrical and optical setups shown in Extended Data Fig. 1 are used for the counting and time-correlated counting in Fig. 3, and the 3D scanning and imaging in Fig. 4. All experiments are done in a standard laser lab environment at room temperature. A power supply provides a d.c. voltage to the d.c. input port of a bias-Tee (Anritsu K251). A pulse function generator (Keysight 81150 A) provides a RF voltage pulse to the RF input port of the bias-Tee and triggers the 1,310 nm pulsed laser (Picoquant LDH with Taiko PDL driver) with a programmable time delay. The output of the bias-Tee is connected to a custom-made RF probe in which alternating-current (a.c.) bypass and 50 Ω termination are implemented. A low-noise amplifier (Mini-Circuits ZFL-1000LN+) is used to amplify the voltage signal to be fed to the counter (Picoquant Multiharp 150) or the oscilloscope (Keysight EXR254A). The output of the pulsed laser is 1% split for optical power monitoring and 99% attenuated by a variable optical attenuator (Thorlabs V1000PA). The output of the variable optical attenuator is 10% split for optical power monitoring and 90% collimated by a collimator. The collimated laser pulses are injected into a custom-made confocal microscope, which consists of a pellicle beam splitter and an objective lens to, respectively, reflect and focus the laser pulses towards the device, and another pellicle beam splitter and two tube lenses to visually monitor the device with two cameras, one visible and the other SWIR. When performing the counting using the counter (that is, for count-rate-related measurements), the event pulse edge and the gate pulse edge are used to trigger and sync the counter, respectively. The counter integration time is typically set to 100 s. When performing the counting using the oscilloscope (that is, for jitter-related measurements), the rising edge of the event pulse is used to trigger the oscilloscope data-capturing frame. Histograms are generated using the rising edge of the event pulse with reference to the gate pulse over 60 K events. When performing the time-correlated counting using the counter (that is, for after-pulsing-related measurements), the output of the low-noise amplifier is divided into two channels with equal power by an RF power divider (custom-made), and a small time delay because of a longer RF cable is introduced to one of the two channels. The outputs of the more delayed channel and the other less delayed channel are, respectively, coupled to the sync and trigger input ports of the counter to enable the time-correlated counting of interarrival time. The counter integration time is typically set to 200 s.

When performing the 3D scanning and imaging using the counter, the variable optical attenuator is replaced by a coaxial transceiver consisting of a circulator, a collimator and x–y galvo-mirrors (Thorlabs QS30XY-AG). Note that a 1,310-nm continuous wave laser and a 635 nm continuous wave laser are coupled to the same optical path of the pulsed laser, and, respectively, used to align the focused laser pulses with the device under the confocal microscope and determine the initial angular direction of steering the mirrors. The laser pulse width and peak power are estimated to be 30 ps and 5 mW, respectively. The mirrors are angular-direction-controlled between x and y angles equal to (−6°, 4°) to (9°, −5°) with 0.1° resolution in both directions in a feedback loop such that the laser pulses are fired according to the angular directions following the steering commands. At every steering command, the acquired histogram from the counter is analysed and the temporal location of its peak is identified and translated into the corresponding x, y and z positions in the PCL dataset. T_g and T_r are set to be equal to 5 ns and 250 ns, respectively, with sufficient integration time for multi-shot LIDAR operation to overcome the high optical loss associated with the whole setup. As T_g is only 5 ns corresponding to the maximum TOF range limited to 75 cm, a space that spans roughly 0.5 m^3 is prepared and filled with a series of 3D objects for photo shoot, in which the centre of the space is located away from the mirrors by roughly 1.15 m.

Finally, the repeatability of DCR and TCR measurements done over 10 consecutive experiments using the counter is determined to be less than 0.1% error (standard deviation divided by mean) for both die-level testing and wafer-level testing; the accuracy of DCR and TCR measurements done at different times is estimated to be about 10% error for die-level testing and less than 10% for wafer-level testing, possibly relating to the fluctuation of ambient temperatures.

Calculation

By fitting $n_{AP}(i)$ with an exponential probability density function $\gamma e^{-t/\tau}$, τ in Fig. 3h can be calculated as

$$\tau = \frac{T_r}{\ln\left(\frac{n_{AP}(1)}{n_{AP}(2)}\right)},$$

and APPs shown in Fig. 3i can be calculated as

$$\mathrm{APP}(i) = \frac{n_{AP}(i)}{1+\sum_{i=1}^{\infty} n_{TG}(i)} \text{ and } \mathrm{APP}^t = \frac{n_{AP}}{(1+\sum_{i=1}^{\infty} n_{TG}(i))\cdot\frac{T_r}{T_g}},$$

where

$$n_{AP}(i) = \int_{iT_r}^{iT_r+T_g} \gamma e^{-t/\tau} dt \text{ and } n_{AP} = \int_0^{\infty} \gamma e^{-t/\tau} dt.$$

Here the interpretation of APP(i) is, under gated-mode operation, the probability of finding an after-pulsing event during the ith gate time when a dark count occurs during the reference gate time. The interpretation of APP^t is, under gated-mode operation, the total probability of finding an after-pulsing event when a dark count occurs.

Data availability

The data supporting this work are available in the main text, figures and Methods. Further information is available from the corresponding author.

Code availability

The codes supporting this work are for extracting the data from the equipment and plotting the data in figures. Further information is available from the corresponding author.

Acknowledgements We thank C. H. Shen and D. Y. Lai for the discussion on the fabrication equipment, process and metrology.

Author contributions N.N. conceived the design, supervised the project, and wrote the paper. Y.-C. Lu contributed to device simulation, layout synthesis and wafer-level data analysis. Y.-H.L. contributed to layout synthesis, process integration and wafer-level data analysis. P.-W.C. and Y.-C. Lai contributed to process integration and device fabrication. Y.-R.L. contributed to process integration and wafer-level data analysis. C.-C.L. contributed to wafer-level data collection and analysis. T.S. contributed to die-level data collection and analysis and 3D PCL image capture. C.-H.C. contributed to die-level data collection and analysis. S.-L.C. oversaw the project.

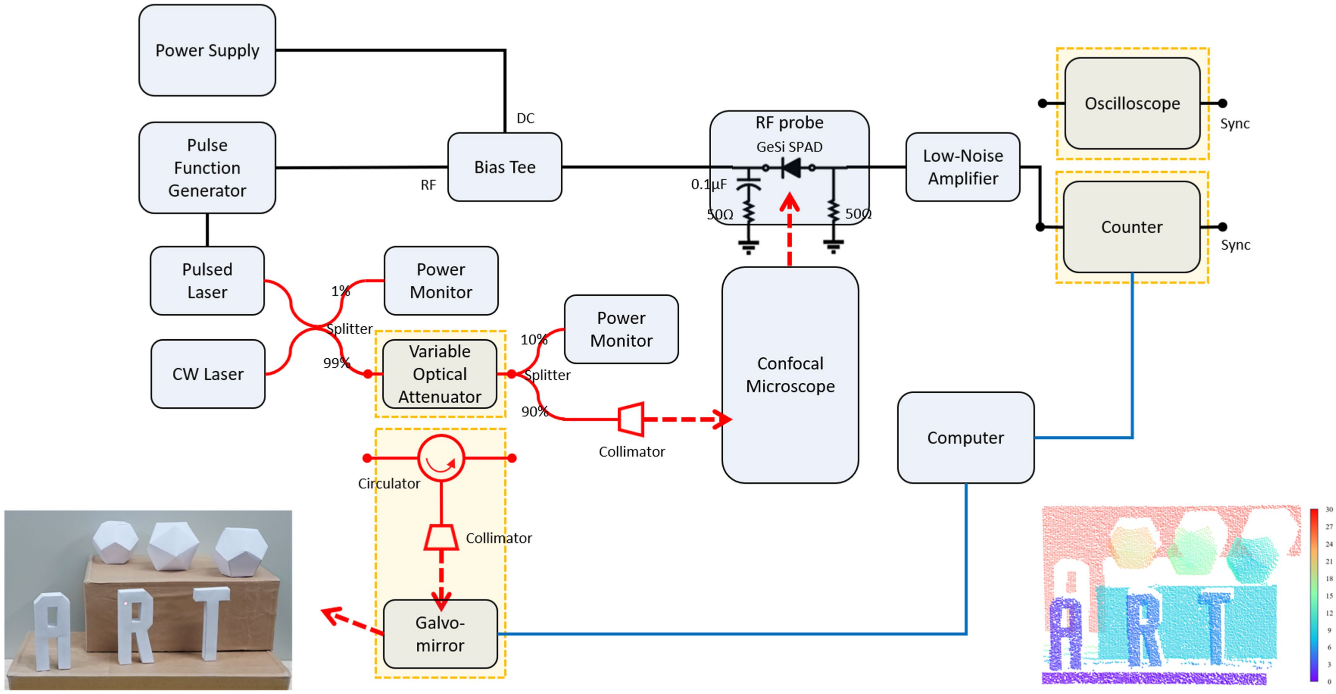
Competing interests All authors are shareholders of Artilux Inc., a start-up company that makes SWIR 2D/3D sensors, imagers and photonic integrated circuits.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-024-07076-x.

Correspondence and requests for materials should be addressed to Neil Na.

Peer review information *Nature* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.



Extended Data Fig. 1 | The schematic plot of the electrical and optical setups used in producing data shown in Figs. 3 and 4.