Sascha Johannes Mühlbrandt

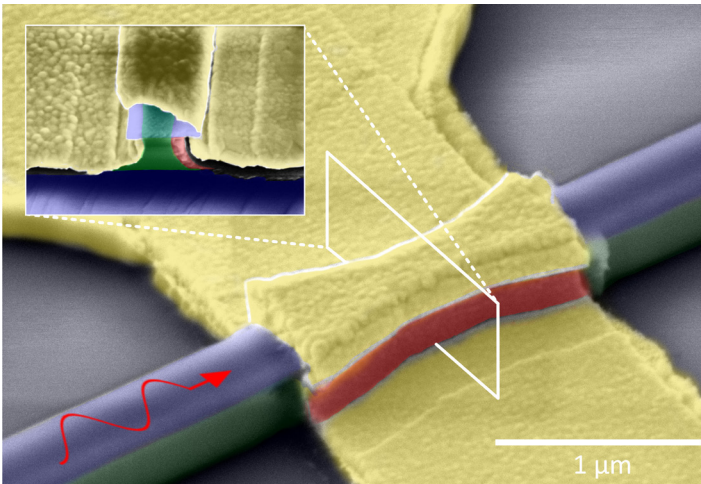# Silicon-Plasmonic Photodetectors for High-Speed Signal Processing

1 µm

KIT Scientific Publishing

Sascha Johannes Mühlbrandt

**Silicon-Plasmonic Photodetectors
for High-Speed Signal Processing**

Silicon-Plasmonic Photodetectors for High-Speed Signal Processing

by
Sascha Johannes Mühlbrandt

Karlsruher Institut für Technologie
Institut für Photonik und Quantenelektronik

Silicon-Plasmonic Photodetectors for High-Speed Signal Processing

Zur Erlangung des akademischen Grades eines Doktors der Ingenieur-
wissenschaften (Dr.-Ing.) von der KIT-Fakultät für Elektrotechnik und
Informationstechnik des Karlsruher Instituts für Technologie (KIT)
genehmigte Dissertation

von Dipl.-Phys. Sascha Johannes Mühlbrandt

Tag der mündlichen Prüfung: 6 Juni 2023
Hauptreferent: Prof. Dr. Christian Koos
Korreferenten: Prof. Dr. Dr. h. c. Wolfgang Freude
                 Prof. Dr. Jasmin Aghassi-Hagmann

Table of Contents

Kurzfassung

Gegenstand der vorliegenden Arbeit ist ein neuartiges Photodetektionskonzept basierend auf interner Photoemission (IPE). Das zugehörige Bauteil wird kurz PIPED genannt, ein Akronym für **Plasmonic Internal Photoemission Detector**. PIPED erreichen als allererste Detektoren aus der Klasse der IPE-Detektoren technisch interessante Leistungskennzahlen. Die Detektoren benötigen eine Fläche von weniger als $1\ \mu m^2$ eines Silizium Chips, erreichen eine optoelektronische Bandbreite von mehr als 1 THz und weisen auch bei diesen hohen Betriebsfrequenzen eine hohe Empfindlichkeit auf. Diese herausragende Leistungsfähigkeit beruht auf dem besonderen Aufbau, der einen Metall-Halbleiter-Metall (MHM) Übergang mit einem Metall-Oxid-Halbleiter (MOH) Übergang verbindet. Aufgrund der obenstehenden einzigartigen Eigenschaften haben Detektoren vom PIPED-Typ das Potential, Schlüsselbausteine zukünftiger photonisch-elektronischer Signalverarbeitungs-systeme mit Bandbreiten im THz Bereich zu werden, die Signale mit bisher unerreichter Geschwindigkeit verarbeiten.

In dieser Arbeit betrachten wir die technische Tragfähigkeit des PIPED Konzepts hinsichtlich solcher photonisch-elektronischer Signalverarbeitungs-systeme. Wir beginnen unsere Untersuchungen mit Anwendungen von PIPED in Chip-zu-Chip Verbindungen. Außerdem untersuchen wir die Eignung und die Limitierungen von PIPED hinsichtlich fortgeschrittener optoelektronischer Signalprozessierung bei sub-THz und THz Frequenzen. Dazu erarbeiten wir ein tiefes Verständnis der grundlegenden physikalischen Prozesse, die die überragende Leistungsfähigkeit des Detektors erst ermöglichen. Zur Herstellung des Bauteils nutzen wir modernste Techniken aus der Halbleiterfertigung.

Die Errungenschaften unserer Forschungsarbeit zum PIPED Konzept wurden mit dem *Gips-Schüle-Forschungspreis 2017* ausgezeichnet, der interdisziplinäre Forschung mit besonderer Relevanz für die Gesellschaft würdigt.

Nachstehend folgt eine detaillierte Beschreibung des Inhalts dieser Arbeit:

Kapitel 1 dient zur Einführung und legt die grundsätzlichen Fragestellungen dieser Arbeit dar.

Kapitel 2 („*Fundamentals of Metal-Dielectric Interfaces*") beschreibt die für das PIPED Konzept wesentlichen theoretischen Grundlagen. Wir stellen insbesondere die physikalischen Prozesse heraus, welche die Kerneigenschaften des PIPED bestimmen.

Kapitel 3 („*State of the Art in Photodetection*") gibt einen Überblick über die Funktionsweise von verschiedenen Photodetektoren. Wir legen hier die Vergleichstechnologien fest, an denen wir die Leistungsfähigkeit des PIPED messen.

Kapitel 4 („*High-Speed Plasmonic Photodetectors*") führt das PIPED Konzept ein. Wir beschreiben den allerersten „Plasmonic Internal Photoemission Detector" (PIPED), der eine hohe Geschwindigkeit und eine Empfindlichkeit von mehr als 0.12 A/W bei einer optischen Eingangswellenlänge von 1550 nm erreicht. Durch Wahl der Polarisation des einfallenden Lichts kann diese Empfindlichkeit noch verdoppelt werden. Wir erreichen mit diesen Bauteilen im Experiment eine Datenrate von 40 Gbit/s, was nur durch die zur Verfügung stehende Laborausrüstung begrenzt wurde. Die funktionsrelevanten Strukturen des Detektors weisen einen Flächenbedarf von weniger als 5 µm × 155 nm auf. Die Gesamtfläche des Detektors ist damit geringer als $1\,\mu m^2$. Diese Leistungskennzahlen stellen Rekordergebnisse für plasmonische Detektoren auf Basis von IPE dar.

Kapitel 5 („*Field-Effect Silicon Plasmonic Photodetector for Coherent T-wave Reception*") erweitert das PIPED Konzept um eine zusätzliche Kontrollelektrode, die in Analogie zu einem MOH-Feldeffekttransistor Gate-Steuerelektrode genannt wird. Durch Anlegen einer Gatespannung wird ein Feldeffekt ausgelöst, mit dem sich die Ladungsträgerdichte im Bauteil kontrollieren lässt. Wir zeigen, dass durch Nutzung dieses Feldeffekts optoelektronische Bandbreiten von mehr als 1 THz möglich werden.

Kapitel 6 („*Silicon–Plasmonic Integrated Circuits for THz Signal Generation and Coherent Detection*") zeigt, dass PIPED sowohl als Transmitter als auch

als Empfänger von T-Wellen dienen können, und sich daher insbesondere für Anwendungen in THz-Systemen eignen. Wir stellen einen Demonstrator vor, der mehrere PIPED- und THz-Schaltkreise monolithisch integriert.

Kapitel 7 („*Summary and Outlook*") umreißt die Kernergebnisse dieser Arbeit und stellt die nächsten Schritte in der Entwicklung und Anwendung zukünftiger Bauteile vom PIPED-Typ hinsichtlich photonisch-elektronischer THz-Signalerzeugung dar.

In den **Anhängen A** bis **D** stellen wir die Techniken zur Herstellung und zur Charakterisierung der PIPED vor. Wir zeigen weitere Messungen, mit denen wir unsere Modelle der grundsätzlichen physikalischen Eigenschaften der PIPED validieren. Zuletzt beschreiben wir die Simulationsumgebungen, mit denen wir die photonischen und elektronischen Eigenschaften unserer Bauteile untersuchen.

Preface

This thesis introduces and experimentally demonstrates a novel approach to exploit internal photoemission (IPE) for photodetection. For brevity, we name the resulting device a PIPED, an acronym for **Plasmonic Internal Photoemission Detector**. PIPED represents the first detector based on IPE that operates at technically viable performance figures. The detectors occupy less than $1\,\mu m^2$ of space on a silicon chip and provide operating frequencies in excess of 1 THz, while a viable sensitivity is maintained even at these speeds. The excelling performance is due to the detector's specific layout that combines a metal-semiconductor-metal (MSM) and a metal-oxide-semiconductor (MOS) junction in a layout similar to a MOS-field-effect transistor. Due to these unique characteristics, PIPED have the potential to become key building blocks of advanced optoelectronic systems that operate at THz frequencies and that process signals at unprecedented speeds.

In this work, we evaluate the technical viability of the PIPED concept regarding such high-speed optoelectronic systems. We begin our analysis with chip-to-chip interconnects based on PIPED. Furthermore, we explore the capabilities and limitations of PIPED for its use in advanced optoelectronic signal processing systems operating in the sub-THz and THz frequency range. We develop a detailed physical understanding of the intrinsic physical processes that define the excelling performance of the detector. We develop a straightforward fabrication process that makes use of state-of-the-art semiconductor processing techniques.

Our achievements regarding the development of PIPED have been awarded with the *Gips-Schüle Research Prize 2017*, which honors interdisciplinary research with distinct benefits for society.

A detailed description of the content follows below:

Chapter 1 ("*Introduction*") provides the general problem statement of this thesis.

Chapter 2 ("*Fundamentals of Metal-Dielectric Interfaces*") introduces the theoretical framework upon which the PIPED concept operates. We explain the physical processes that define the key characteristics of PIPED.

Chapter 3 ("*State of the Art in p-i-n* Photodetect") describes the general operating principle of photodetectors. Here, we define the benchmark technologies with which the PIPED compete.

Chapter 4 ("*High-Speed Plasmonic Photodetectors*") introduces the PIPED concept. We demonstrate the first high-speed and high-sensitivity plasmonic internal photoemission detector (PIPED) with a sensitivity exceeding 0.12 A/W at 1550 nm. By changing the state of polarization, the sensitivity can be doubled. We succeed in data reception at bit rates up to 40 Gbit/s using a PIPED receiver, where the data rate was only limited by the equipment available for the experiment. The functionally relevant structures occupy an area of only 5 µm × 155 nm. The total footprint of the device is less than 1 µm^2. Our experimental findings represent record results for plasmonic photodetectors based on IPE.

Chapter 5 ("*Field-Effect Silicon Plasmonic Photodetector for Coherent T-wave Reception*") expands the PIPED concept by introducing a further control electrode, which in analogy to a MOSFET is called the gate electrode. Applying a voltage to the gate introduces a field effect that controls the charge carrier distribution and the associated dynamics in the device. We show that optoelectronic bandwidths in excess of 1 THz are feasible using the gate.

Chapter 6 ("*Silicon–Plasmonic Integrated Circuits for THz Signal Generation and Coherent Detection*") shows that PIPED can operate as both, T-wave transmitters and receivers, and hence lend themselves for application in T-wave systems. We suggest a proof-of-concept design, in which we monolithically integrate PIPED with T-wave circuits.

Chapter 7 ("*Summary and Outlook*") formulates the key findings of this thesis and details the next steps in the development and the application of future PIPED-type devices in optoelectronic signal processing systems operating at THz frequencies.

In the **appendices A** through **D**, we describe the technologies we use to create and experimentally analyze the PIPED. We supplement further measurement data that we use to validate our models of the physical properties of the PIPED. Lastly, we explain the simulation framework with which we model the photonic and electronic characteristics of our devices.

Achievements of the Present Work

This thesis introduces and experimentally demonstrates a novel silicon plasmonic photodetection concept. It is based on internal photoemission (IPE), which is the light-induced transfer of hot carriers across a metal-dielectric interface. We name the resulting device a PIPED, an acronym for **Plasmonic Internal Photoemission Detector**. In this work, we develop the PIPED concept starting with the analysis of its fundamental physical operating principles, and investigate possible application scenarios for PIPED, ranging from chip-level interconnects to optoelectronic terahertz wave (T-wave) systems.

The main achievements of this work are listed below:

First demonstration of a silicon plasmonic photodetector with technically interesting performance: We experimentally demonstrate a PIPED with a record-high sensitivity of more than 0.12 A/W at 1550 nm and with large optoelectronic bandwidths well above 40 GHz. We prove the viability of the PIPED concept for real-world applications by receiving on-off keying (OOK) data at a rate of 40 Gbit/s. Our experimental findings are only limited by the available equipment at the time of the experiment.

Development and implementation of a straightforward PIPED nano-fabrication scheme in silicon: Our fabrication process is able to create PIPED with a device footprint of below 1 μm^2. This makes PIPED much smaller than competing Si/Ge photodetectors or uni-travelling-carrier (UTC) photodiodes, which typically consume up to 100 μm^2.

Ultra-wide range of operation: The optical operating frequency $f_0 = c/\lambda_0$ is only limited by the absorption band edge of silicon at high optical frequencies and by the barrier height of the involved metals at low optical frequencies. For our PIPED design, we expect a useful operating range from $\lambda_0 = 1130$ nm to above 2000 nm. We have yet to determine the optimal operating wavelength that maximizes the sensitivity.

Ultra-high optoelectronic bandwidth > 1 THz: Employing a field-effect allows to control the carrier dynamics in the PIPED. We show that this field-

effect is persistent even at T-wave frequencies. Our simulations show that this field-effect can be utilized to increase the optoelectronic bandwidth to well above 1 THz.

First demonstration of silicon plasmonic T-wave transmitters and receivers: Exploiting the high bandwidth and the voltage-dependent sensitivity of PIPED, a single device can both generate T-waves by photomixing, and receive T-waves by optoelectronic down-conversion using an optically generated local oscillator.

First demonstration of integrated T-wave systems: We show that PIPED can be monolithically cointegrated with T-wave circuits. We determine the complex transfer characteristics of the transmission line in the frequency range from 0.05 THz to 0.8 THz as a proof-of-concept experiment.

1   Introduction

The demand for high-performance computing and ultra-broadband signal processing is ever increasing. Advanced optoelectronic systems attempt to address this demand by cointegration of photonic and electronic elements [1, 2]. This allows to make use of the most advantageous characteristics of both fields: Integrated electronic circuits process digital information by relying on strongly nonlinear components (like field-effect transistors). State-of-the-art electronics achieve ultra-high integration densities. As of today, the fabrication techniques in semiconductor structuring have reached a level of precision that allows to control structures nearly down to atomic levels. Photonics, on the other hand, has spawned a multitude of signal processing techniques that provide broadband tunability and ultra-wide bandwidths. High-performance computing systems that merge the strengths of both fields could operate at very high frequencies and offer unsurpassed performance metrics. High-speed photodetectors are core elements of such optoelectronic systems. They convert light to an electronic current and hence mediate between the photonic and the electronic domain.

However, a seamless cointegration of photodetectors and state-of-the-art electronics is currently impeded by vastly different boundary conditions imposed by the underlying technologies: Whereas modern field-effect transistors can be fabricated with feature sizes approaching the single-digit nanometer range [3], typical dimensions of integrated photodetectors amount to several micrometers. Furthermore, while modern CMOS electronics are fabricated on the mature silicon platform, photodetectors are typically fabricated in germanium or on dedicated III–V substrates, both of which are not readily transferable to highly advanced CMOS platforms. This impedes a monolithic cointegration of high-speed photodetectors in silicon electronics.

In this work, we introduce **Silicon Plasmonic Internal Photoemission Detectors** (PIPED) as a new class of ultra-compact high-speed photodetectors. PIPED rely on the excitation of charge density oscillations at metal-dielectric interfaces, so called surface plasmon polaritons (SPP), and on internal

photoemission (IPE), i.e., the light-induced transfer of "hot" carriers across such interfaces.

PIPED implement SPP and IPE in a quasi-ideal way: Light is coupled from a silicon photonic waveguide to a metal-semiconductor-metal (MSM) junction made of gold, silicon, and titanium. The silicon core is less than 100 nm wide, which creates a high plasmonic field enhancement at the junction and leads to a strong SPP absorption. This generates electron-hole pairs in the metals within a nanometric distance from the interface – close enough to allow a hot carrier transfer across the metal-semiconductor interfaces, thereby creating a photocurrent.

In the following chapters, we explore and develop the PIPED concept in multiple steps: Starting from its very fundamentals, we create a detailed understanding of the device's underlying physical processes. We implement a fabrication scheme to create PIPED with less than 1 µm² in footprint. Specifically, our PIPED is the first IPE-based photodetector that shows technically viable performance parameters, which are comparable to other state-of-the-art high-speed photodiodes. For the first time, high-speed photodetection at 1550 nm wavelength becomes possible in silicon. With these early PIPED prototypes, we are already able to demonstrate record-high sensitivities in excess of 0.12 A/W at bandwidths of more than 40 GHz [4]. This makes PIPED a worthwhile candidate for ultra-compact and highly integrated chip-to-chip interconnects.

We then extend the PIPED concept by introducing a metal-oxide-semiconductor (MOS) interface with an additional gate electrode [5]. Applying a voltage to the gate induces a field effect in the silicon device core, similar to a MOS field-effect transistor (FET). This allows to control the carrier dynamics in the device and the degree of internal photoemission at the metal-semiconductor interfaces. Specifically, a highly conductive channel can be created in the device in which the device dynamics are determined by ultra-fast dielectric relaxation rather than by the carrier transit time. We experimentally verify that this field effect is persistent even at sub-THz frequencies. Our simulations show that the bandwidth of PIPED can be increased to more than 1 THz by using the gate.

Furthermore, we analyze the capabilities and limitations of PIPED for its use in advanced optoelectronic signal processing systems that operate in the sub-THz and THz frequency range [6]: PIPED feature a strongly voltage-dependent sensitivity and are, thus, not only suited for T-wave generation by photomixing, but also for coherent detection of T-waves by optoelectronic down-conversion when employing an optical local oscillator (LO). PIPED can hence act both as T-wave transmitters and as receivers. We demonstrate a monolithically integrated T-wave system based on PIPED on the silicon platform and measure the complex-valued transfer characteristics of an on-chip transmission line as a proof of principle. We create T-waves in a PIPED transmitter, and continuously tune the frequency range from 0.05 THz to 0.8 THz to extract the amplitude and phase information at the PIPED T-wave receiver.

We believe that the PIPED concept addresses multiple challenges that are faced by optoelectronic signal processing systems. These challenges range from fundamental device principles to system integration. PIPED have the potential to become disruptive in a variety of applications, comprising, without limitation, future chip-level interconnects [1], ultra-broadband wireless communications in future sixth-generation (6G) networks [7, 8], or THz spectroscopy systems [9].

2 Fundamentals of Metal-Dielectric Interfaces

This chapter introduces the theoretical framework describing the material properties and the physical processes upon which this work is built. Section 2.1 describes the basic electronic properties of solids, introducing the continuity equation and discussing carrier-transport mechanisms in semiconductors. Section 2.2 focuses on interfaces of metals and semiconductors and on interfaces of metals, oxides, and semiconductors, respectively. The carrier transport under non-equilibrium conditions is specifically in focus. Section 2.3 introduces surface plasmon polaritons (SPP), which are a special form of electro-magnetic waves bound to metal-dielectric interfaces.

2.1 Electronic Properties of Solids

The following section provides a brief overview of the fundamental physical properties of semiconductors. A detailed derivation of the findings from first principles can be found in [10] and [11].

Band Structure

The electronic properties of matter in the solid state are described with the energy band model. The energy bands are a consequence of the wave-like nature of electrons and their interaction with the periodic crystal lattice. Due to the periodicity of the crystal lattice, only distinct waves with propagation constant k and associated energy states $W(k)$ can exist. The crystal momentum $\hbar k$ is akin to the classical momentum and is a conserved quantity. Each occupied state models an electron. An unoccupied state is referred to as a "hole".

The bands determine the electronic states of energy $W(k)$ that electrons can occupy and are separated by band gaps, where no states exist and electrons cannot dwell. The band gap energy is usually measured in units of electron volt and is typically in the order of $W_G = 1$ eV in semiconductors (see Table 1). Insulators have a larger band gap of several eV. For metals, the bands overlap and there is no band gap.

Table 1. Band gaps for various materials.

Material	Type	Band Gap W_G
Silicon	Semiconductor	1.1 eV
Germanium	Semiconductor	0.66 eV
Silicon oxide	Insulator	~10 eV

Effective Mass and Mobility

Free charge carriers in the solid are able to move under the influence of an external force. The effective masses of electrons m_n^* in the conduction band and of unoccupied electron states ("holes") m_p^* in the valence band are a measure of the carrier movement in response to an external field, in analogy to the inertia of an object with mass m in classical mechanics. The effective masses are characteristic for each material and are strongly dependent of the energetic state $W(k)$ in the k-space that the electron (or hole) occupies. The effective mass reads in one dimension

$$\frac{1}{m^*} = \frac{1}{\hbar^2} \frac{\partial^2 W(k)}{\partial k^2} . \tag{1}$$

The effective masses of electrons and holes are typically different due to the different curvatures of the band edges at the bottom of the conduction band and at the top of the valence band. The effective mass can be assumed constant in most practical applications, as scattering limits the amount of momentum Δk that a carrier can accumulate, such that the parabolic approximation according to Eq. (1) is valid and the effective mass m^* can be considered as a constant. This scattering finally limits the speed of the carriers, which move with the group velocity of the associated wave packet

$$v_g = \frac{1}{\hbar} \frac{\partial W(k)}{\partial k} . \tag{2}$$

In the following equations, only the x-dimension is assumed for the sake of simplicity. The derivation is valid for the y- and the z-dimension, too.

The mobility μ relates this average speed v_x and the electric field E_x and takes into account the scattering time τ_{SC} of carriers when interacting with impurities or phonons:

$$|v_x| = \mu |E_x| = \frac{e\tau_{SC}}{m^*} |E_x|. \tag{3}$$

Occupation Probability and Density of States

The Fermi-Dirac distribution function $f(W)$ determines the occupation probability of any electron state at energy W. For states within the bands, it reads

$$f(W) = \frac{1}{1 + \exp\left[(W - W_F)/(kT)\right]} \overset{W - W_F \gg kT}{\approx} \exp\left[-(W - W_F)/(kT)\right], \quad (4)$$

where W_F is the Fermi energy, k is the Boltzmann constant and T the absolute temperature. At room temperature ($T = 298$ K) the Fermi function makes a transition from high (0.88) to low (0.12) occupation probabilities within an energetic width of $4kT = 100$ meV.

Conversely, the probability that a state is unoccupied is given by

$$1 - f(W) = \frac{1}{1 - \exp\left[-(W_F - W)/(kT)\right]} \overset{W_F - W \gg kT}{\approx} 1 - \exp\left[-(W - W_F)/(kT)\right]. (5)$$

In semiconductors, the highest occupied band is called the valence band and the highest unoccupied band is the conduction band. In a three-dimensional system, the density of states follows a square-root law close to the bottom of the conduction band N_C and the top of the valence band N_V (page 17 in [10])

$$N_C(W) = \frac{m_n^* \sqrt{2m_n^*(W - W_C)}}{\pi^2 \hbar^3}, \tag{6}$$

$$N_V(W) = \frac{m_p^* \sqrt{2m_p^*(W_V - W)}}{\pi^2 \hbar^3}. \tag{7}$$

Carrier Density and Doping

The intrinsic carrier density n_i is the number of electrons in the conduction band, and the number of holes in the valence band which are present due to thermal

excitation. For semiconductors, n_i is typically small of the order of 10^{10} cm^{-3} due to the large energetic separation of the bands. The carrier densities in the semiconductor can be greatly increased by doping the material with an acceptor (density N_A) or with a donor (density N_D). Acceptors introduce additional energy states close to the edge of the valence band. Electrons from the valence band can be thermally excited to occupy an acceptor state, leaving a hole and an ionized acceptor behind. Similarly, donors introduce additional energy states close to the band edge of the conduction band. Electrons from donor states can be thermally excited to the conduction band. The energetic separation of donors and acceptors from the band edges is typically of the order of 10 to 100 meV and hence much smaller than the band gap energy. In thermal equilibrium and at room temperature, all acceptors and donors are ionized, such that the density of carriers in the semiconductor is

$$n_0 = N_D = N_C \exp\left(-\frac{W_C - W_F}{kT}\right), \tag{8}$$

$$p_0 = N_A = N_V \exp\left(-\frac{W_F - W_V}{kT}\right). \tag{9}$$

The carrier density can be manipulated by a potential $\varphi(x)$ that is present in the semiconductor. The potential $\varphi(x)$ can be the result of a contact voltage at material interfaces, or it can be due to an externally applied voltage. The reference potential $\varphi = 0$ is located outside of the disturbed domain, where the semiconductor is in thermal equilibrium. The carrier density then reads:

$$n = n_0 \exp\left(-\frac{e\varphi(x)}{kT}\right), \tag{10}$$

$$p = p_0 \exp\left(\frac{e\varphi(x)}{kT}\right). \tag{11}$$

The dominating dopant type determines the majority carrier species. The respective other carrier type is called the minority species. The mass action law states the relation between n, p and n_i in thermal equilibrium:

$$n_i^2 = np \tag{12}$$

Conductivity and Carrier Transport

Transport is the physical movement of carriers through the solid at speed v_x, forming an electric current J. The current is counted positive for positively charged carriers travelling along the x-direction (positive v_x) and negative in case of negatively charged carriers. Transport can occur due to gradients in carrier densities or due to external fields. Gradients lead to diffusion currents J_D, which are determined by the diffusion constant $D = kT\mu/e$. The relations for diffusion currents of electrons and holes read:

$$J_{n,D,x} = eD_n \frac{dn}{dx} \quad , \tag{13}$$

$$J_{p,D,x} = -eD_p \frac{dp}{dx}. \tag{14}$$

The drift current $J_{F,x}$ due to an external field E_x can be derived from Eq. (3), and by using the relation for the conductivity for electrons $\sigma_n = en\mu_n$ and $\sigma_n > 0, E_x > 0$:

$$J_{n,F,x} = -env_{n,x} = -en\mu_n E_x = -\sigma_n E_x. \tag{15}$$

An analogous expression holds for holes:

$$J_{p,F,x} = epv_{p,x} = ep\mu_p E_x = \sigma_p E_x. \tag{16}$$

Generation and Recombination

Electron-hole pairs are constantly generated and recombined in semiconductors. In thermal equilibrium and without any externally applied voltages or illumination, the generation rate g and the recombination rate r cancel each other, i.e., they have the same magnitude and opposite signs. The net generation or recombination rate for carriers can be influenced by:

- Heating of the semiconductor as a whole by increasing the temperature T
- Light absorption in the semiconductor, causing inter-band transitions
- Injection of charges from the outside (e.g., through an interface to a different material)

Continuity Equation

The continuity equation relates the spatial distribution of free carriers in x-, y-, and z-direction and the associated temporal evolution to the associated currents as well as to the generation and recombination of the carriers in the semiconductor. The equation of continuity is a result of the conservation of mass and electric charge: A given density of charge carriers can only change in space and time if there is a net inflow or outflow of carriers, or if the carriers are generated or recombined through transitions between the valence band, the conduction band and the impurity states within the band gap.

The continuity equation for electrons reads:

$$\frac{\partial n}{\partial t} - \frac{1}{e}\mathrm{div}\vec{J}_{\mathrm{n}} = (g_{\mathrm{n}} - r_{\mathrm{n}}), \quad \vec{J}_{\mathrm{n}} = \vec{J}_{\mathrm{n,F}} + \vec{J}_{\mathrm{n,D}} = en\mu_{\mathrm{n}}\vec{E} + eD_{\mathrm{n}}\mathrm{grad}(n). \quad (17)$$

An analogous relation holds for holes:

$$\frac{\partial p}{\partial t} + \frac{1}{e}\mathrm{div}\vec{J}_{\mathrm{p}} = (g_{\mathrm{p}} - r_{\mathrm{p}}), \quad \vec{J}_{\mathrm{p}} = \vec{J}_{\mathrm{p,F}} + \vec{J}_{\mathrm{p,D}} = ep\mu_{\mathrm{p}}\vec{E} - eD_{\mathrm{p}}\mathrm{grad}(p). \quad (18)$$

Dielectric Relaxation

Dielectric relaxation is a mechanism of carrier transport in semiconductors occurring in domains of high carrier density. Injection of minority or majority carriers into an otherwise electrically neutral region leads to a disturbance of the potential $\varphi(x)$. The dominating carrier species reacts to the disturbance by reorganizing its distribution and by electrically shielding the excess charge. The high electrical conductivity of the majority species allows for a rapid redistribution of the carrier density. The dielectric relaxation time τ_{R} is the characteristic time that a material with a given conductivity σ needs to shield excess charges locally and is given by (page 519 in [10])

$$\tau_{\mathrm{R}} = \frac{\varepsilon}{\sigma}. \quad (19)$$

Typical values for the dielectric relaxation time range from $\tau_{\mathrm{R}} = 0.5\mathrm{fs} \ldots 0.5\mathrm{ps}$, assuming an electron concentration between $n = 10^{16}$ cm^{-3} and 10^{19} cm^{-3} in n-doped silicon.

The bandwidth of optoelectronic devices can benefit strongly from dielectric relaxation. The carrier dynamics for such devices is determined by the usually very small dielectric relaxation time instead of the carrier transit time. In the regime of dielectric relaxation, very high bandwidths in excess of 1 THz can be achieved, see Chapter 5.

2.2 Metal-(Oxide)-Semiconductor Interfaces

This section explains the electronic properties of metal-semiconductor (MS) and metal-oxide-semiconductor (MOS) interfaces. A metal-semiconductor-metal (MSM) junction combines two MS-interfaces, which share the same semiconductor layer. Figure 1 (a) shows an illustration of a combined MSM and MOS structure. The MSM junction is oriented horizontally and consists of a silicon (Si) core between a layer of gold (Au) and a layer of titanium (Ti). In the context of this work, the width of the silicon core layer is of the order of $w = 100$ nm. There is an additional silicon oxide (SiO_2) layer and a gold cap on the top of the MSM junction, which establishes a MOS interface in the vertical direction. The individual metal structures form electrodes and are called *gate* for the gold cap, *drain* for the gold contacting the silicon and *source* for the Ti contact in analogy to a MOS field-effect transistor. The SiO_2 layer below the gate will be called *gate oxide* in the following. The potential $\varphi(x,y)$ in the silicon may be controlled by applying an external voltage U_{DS} between the drain and the source contact or U_{GS} between the gate and the source. The voltages U_{DS} and U_{GS} are counted positively from the drain (or the gate) contact with respect to the source. The following sections explain the key aspects of this combined MSM/MOS junction with respect to the band structure and to the charge transport properties.
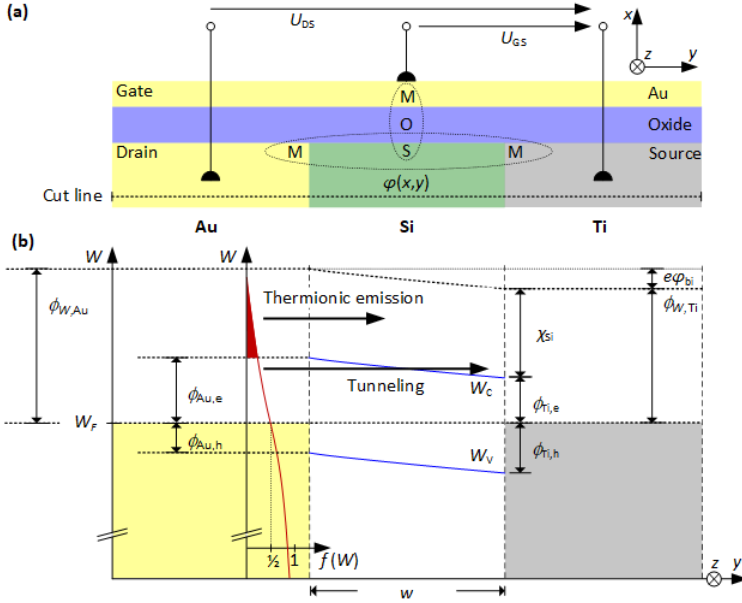
Figure 1. Illustration of a combined metal-semiconductor-metal (MSM) / metal-oxide-semiconductor (MOS) structure and of the associated band diagram. **(a)** The MSM part of the structure consists of a silicon core between a layer of gold (drain contact) and a layer of titanium (source contact) in a horizontal direction. In the vertical direction, there is an additional silicon oxide layer and a gold cap (gate) on the top of the MSM junction, which establishes a metal-oxide-semiconductor (MOS) interface. The voltages U_{DS} and U_{GS} are counted positively from the drain (or the gate) contact with respect to the source. **(b)** Band diagram following the cut line in (a) of the Au-Si-Ti junction without an external voltage. The difference of the metal work functions $\phi_{W,Au}$ and $\phi_{W,Ti}$ gives rise to a built-in potential φ_{bi}. The associated voltage drops across the silicon core of width w. The potential steps at the respective MS interfaces $\phi_{Au/Ti}$ are determined by the difference of the work function ϕ_W and the electron affinity χ_{Si}, as well as the band gap. The Fermi energy W_F is constant throughout the complete junction if no voltage is applied. The slope of the silicon conduction band edge W_C and of the valence band edge W_V is mainly determined by the built-in potential φ_{bi}. Charges in the metals can overcome the barrier either by thermionic emission or by tunneling, depending on their energy. Thermionic emission is the direct transition of carriers over the barrier $\phi_{Au/Ti}$ if they have an energy $W > W_F + \phi_{Au/Ti}$. Tunneling is the classically forbidden transition through the barrier if the energy is $W < W_F + \phi_{Au/Ti}$. This is exemplified for electrons in Au: The electrons in the high-energy tail of the Fermi distribution can directly overcome the barrier (highlighted in red), while the residual electrons can only attempt to tunnel. The emission of holes from the gold and the emission of electrons from the titanium is strongly suppressed without an external bias: The intrinsic field due to the built-in potential works against the carriers that attempt to tunnel through the barrier.

2.2.1   Band Structure

Figure 1 (b) shows the band structure of the MSM junction along a cut line from the Au drain to the Ti source contact and through the silicon core. The trace of the cut line is drawn in Figure 1 (a). The silicon core is weakly p-doped with an acceptor concentration of $N_A \approx 10^{15}$ cm^{-3} to compensate for impurities. The exact shape of the band structure is determined by the Fermi energy level W_F, by the potential steps at the MS interfaces, and by the curvature of the silicon conduction and valence band edges.

The Fermi energy level W_F must be constant across the entire junction in the absence of external voltages. Before contact, the gold and the titanium exhibit different work functions $\phi_{W,Au}$ and $\phi_{W,Ti}$ respectively. Bringing gold, silicon, and titanium in contact leads to an exchange of charges until the Fermi energy is flat across the junction.

The net potential across the MSM junction that is created due to the charge exchange is $\varphi_{bi} = (\phi_{W,Au} - \phi_{W,Ti})/e$. The voltage associated with the built-in potential only drops across the semiconductor. The metals are free of excess charges except for a thin charged layer directly at the respective MS interfaces.

The MS interfaces constitute energy steps in the conduction and the valence bands, which are defined by the differences of the Fermi energy W_F (dotted line) and the band edges of the silicon (blue lines). The steps exist due to the band gap, which prevents electrons and holes from the metals to enter the silicon. In an ideal case, the barrier heights are determined by the difference of metal work function ϕ_W and the silicon electron affinity χ_{Si}. The barrier height for electrons at the Au-Si interface is $\phi_{Au,e} = \phi_W - \chi_{Si}$. The barrier height for holes is $\phi_{Au,h} = \phi_W - (\chi_{Si} + W_{G,Si})$. Note that $\phi_{Au,h}$ has a negative sign because it is counted negatively on the energy scale of electrons (W-axis). Analogous expressions hold for the Si-Ti interface.

The actual barrier heights that can be encountered in real MS junctions typically differ from the ideal case. The barrier heights strongly depend on the materials and on the preparation of the surfaces. Surface states created by the semiconductor processing or gold diffusion might significantly affect the actual barrier height. The assumed barrier heights for the scope of this work are listed

in Table 2. To confirm those assumptions, Appendix B.1 discusses an experimental way to measure the barrier height at the Au-Si interface. Device measurements and numerical analysis of the structure in Appendix C.2 allows to determine the built-in potential $\varphi_{bi.}$ The remaining quantities can be calculated with the above relations. Hence, a full set of information about the energetic properties of the Au-Si-Ti junction can be achieved.

At zero external voltages $U_{DS} = 0$ and $U_{GS} = 0$, and away from the oxide-semiconductor interface, the curvature of the silicon conduction band edge W_C and the valence band edge W_V is mainly determined by the built-in potential φ_{bi}, as the contribution from the free carrier densities $n(y)$ and $p(y)$ and the ionized acceptor concentration N_A^- in the silicon is negligible. We confirmed this assumption by a numerical model of the entire MSM/MOS structure. The following section explains our reasoning.

At the MOS junction, there is no carrier injection due to the zero-gate potential. The silicon core is otherwise completely depleted of free carriers due to the overlapping space-charge regions of both MS-interfaces. Hence, the influence of the charge carrier densities n and p can be neglected. We confirmed this assumption by a numerical model of the entire MSM/MOS structure. The additional curvature of the band edges due to ionized acceptors in the space-charge region can be obtained from considering the charge density $\rho = -eN_A^-$. When assuming a silicon core width $w = 100$ nm, the potential due to the ionized dopants is of the order of 10 mV, which is much smaller than the built-in potential. It is hence fair to assume that the built-in potential dominates the behavior of the band edges. In this case, the built-in potential drops linearly across the silicon core and the electric field is constant. The flat-band condition applies when the band edges are entirely horizontal. This can be achieved by applying the built-in potential between the drain and the source contact $U_{DS} = \varphi_{bi}$. Section 2.2.2 and Chapter 5 discuss the influence of the voltages on the band structure in more detail.

Table 2. Barrier heights at the Au-Si and the Si-Ti interface and the resulting built-in potential φ_{bi}. The data has been confirmed by measurements and numerical simulation.

Gold-silicon interface		Barrier height
Electron barrier	$\phi_{Au,e}$	0.82 eV [12, 13]
Hole barrier	$\phi_{Au,h} = \phi_{Au,e} - W_G$	–0.29 eV
Silicon-titanium interface		Barrier height
Electron barrier	$\phi_{Ti,e}$	0.62 eV [14]
Hole barrier	$\phi_{Ti,h} = \phi_{Ti,e} - W_G$	–0.49 eV
Built-in potential across silicon		Potential
	$\varphi_{bi} = (\phi_{Au,e} - \phi_{Ti,e})/e$	0.2 V

2.2.2 Hot Carrier Transport across Metal-Semiconductor-Metal Junctions

The charge transport across the junction mainly depends on the presence of hot carriers in the metal. Low-energy carriers near the Fermi energy level of the metals cannot directly enter the silicon, as the band gap does not provide free states. The band gap constitutes a current-blocking barrier. However, carriers which feature a large enough energy can cross the silicon core.

To understand the underlying physical processes, it is helpful to approach the topic in two steps. In a first step, analyzing the behavior of individual particles at the MS interface allows to understand how single hot carriers can contribute to a current across the junction. As a second step, analyzing the behavior of the carriers as a statistical ensemble helps to understand how heated carrier "gases" in the metals may contribute to the current. Heating of the carrier gas can occur if hot electron-hole pairs relax and share their energy with the surrounding electronic system. The following sections describe the details.

The transmission probability $T(W)$ of a single electron at energy W attempting to overcome a potential barrier strongly depends on the exact shape of the barrier. For a simplified analysis, a rectangular step barrier with a barrier height ϕ_B and a barrier width w can be considered. We further neglect differences of

the electron effective masses inside and outside of the barrier and set the effective mass m^* equal to the mass m_e of the electron particle in free space. With above simplifications, an analytical solution can be found by solving Schrödinger's equation for the wave function inside and outside of the barrier and by considering the continuity of the wave function at the interfaces of the potential barrier. The solution for the transmission probability is valid for the complete energy range above and below the potential step barrier and is given in one dimension by (page 48 in [10]):

$$T(W) = \left[1 - \frac{\phi_\mathrm{B} \sinh\left(\dfrac{w}{\hbar} \sqrt{2m_\mathrm{e}(\phi_\mathrm{B} - W)} \right)}{4W(\phi_\mathrm{B} - W)} \right]^{-1} \tag{20}$$

The following sections distinguish two different energetic regimes. Tunneling applies if the carrier energy is below the barrier energy. Thermionic emission applies when the carrier energy is above the barrier energy.

Tunneling, $W < W_\mathrm{F} + \phi_\mathrm{B}$

Tunneling is the transition of carriers through energetically forbidden regions of the barrier, i.e., for the case in which the carrier energy level W is smaller than the energy level of the barrier, $W < W_\mathrm{F} + \phi_\mathrm{B}$. Tunneling is possible in the quantum regime due to the wave-like nature of carriers (section 1.5.7 in [10]). The inset of Figure 2 shows an electron wave impinging on a step-like potential barrier. Classically, the electron is forbidden to enter the potential barrier. In a quantum picture, however, the tail of the carrier wave function penetrates into the barrier and exponentially decays along the barrier width. If the barrier is sufficiently small, the probability density of the carrier at the far end of the barrier is still large enough to enable the transition of the carrier. Figure 2 shows the transmission probability $T(W)$ for a rectangular barrier for widths of $w = 1$ nm and $w = 10$ nm according to Equation (20). Although the ratio of the barrier thicknesses is only 1:10, the transmission probability is virtually zero for the case of the thicker barrier, while it is significantly larger for the thin barrier. This demonstrates the high sensitivity of the tunneling process to the barrier width. The transmission probability only becomes significant when the carrier energy W approaches or exceeds the energy level of the barrier $W_\mathrm{F} + \phi_\mathrm{B}$.
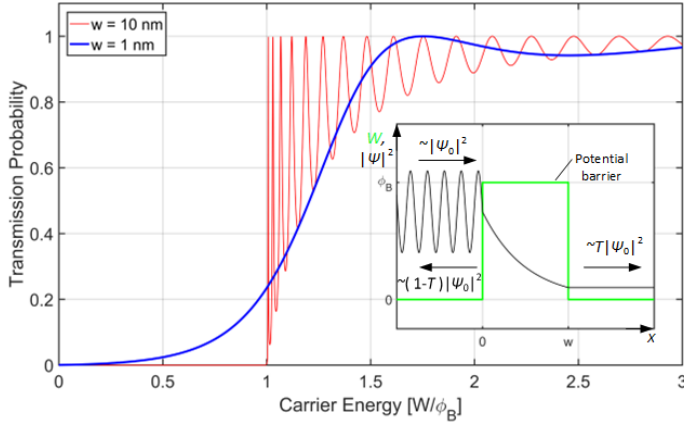
Figure 2. Transmission probability $T(W)$ for an electron at energy W attempting to overcome a rectangular barrier with a barrier height ϕ_B in one dimension. The blue line depicts the probability in the case of a thin barrier of width $w = 1$ nm while the red line refers to the case of $w = 10$ nm. The energy range $W < \phi_B$ is the tunneling domain. Here, the transmission probability is significantly smaller than 1. Although the ratio of the two barrier widths is only 1:10, the transmission probability is virtually zero for the case of the thicker barrier when $W < \phi_B$. For $W > \phi_B$, in the domain of thermionic emission, a direct transition across the barrier without tunneling becomes possible. For these direct transitions, the classically expected probability is one. This is not the case here due to the wave properties of the carriers. The inset shows the transmission of a wave with amplitude Ψ_0 and energy $W < \phi_B$ through a barrier. The particle influx is proportional to $j \sim \Psi^* d\Psi/dx - \Psi d\Psi^*/dx$. The wave enters the barrier and its amplitude $\left|\Psi_0\right|^2$ is exponentially decreasing inside the barrier. The transmission probability is $T < 1$. The wave is partly reflected, leading to a standing wave of the probability density in front of the barrier.

Note that the analysis of Equation (20) constitutes only a simplified view and omits several constraints that influence the transmission probability at real interfaces. First, the effective masses m^* at the MS interfaces are discontinuous, causing additional reflection of waves. Furthermore, the statistical momentum and the energy distribution of the carriers in three dimensions must be considered to determine the number of available carriers for tunneling. Only the fraction of carriers that feature a momentum component towards the potential step can attempt to tunnel. Tunneling is especially relevant for asymmetric structures that exhibit rectifying characteristics under bias, like MS interfaces

17

or Zener diodes. Large biases progressively reduce the effective barrier width. If large enough voltages are applied, the tunneling probabilities can become so large that the rectifying behavior of the junction disappears, and the interface behaves like an ohmic conductor.

Thermionic Emission $W > W_F + \phi_B$

Thermionic emission is the direct transition of high-energy carriers across the barrier, i.e., tunneling is excluded. This is possible if the carrier energy W exceeds the energy level of the barrier at $W_F + \phi_B$ and if the carrier travels in the direction of the barrier. Hence, only hot carriers can enter the semiconductor. Considering a hot carrier gas at a single MS interface at temperature T, the current from the metal into the semiconductor due to thermionic emission is (page 47 in [10])

$$J_{m \to s} = -A^* T^2 \exp\left(-\frac{e\phi_B}{kT}\right). \tag{21}$$

In this relation, A^* is the effective Richardson constant and depends on the semiconductor's effective masses. For a MSM junction in thermal equilibrium and without external voltages $U = 0$, the total thermionic emission currents from all individual MS interfaces and carrier species must be zero, $J_{TE} = \sum\limits_{Au\text{-}Si,\ Si\text{-}Ti} J_{m \to s} = 0$.

Classically, all carriers with $W > W_F + \phi_B$ would overcome the potential barrier. However, this is not the case in the quantum domain. The transmission probability in Figure 2 shows an oscillatory behavior for energies $W > W_F + \phi_B$. This is due to standing waves that emerge in the barrier, similar to the behavior of light in a Fabry-Pérot-etalon.

Applying a voltage $U_{DS} \neq 0$ to an MSM junction allows to control the relative thermionic emission current at each MS interface (note: The gate voltage is $U_{GS} = 0$ and there is no carrier injection from the MOS junction). Figure 3 shows the band structure of an MSM junction with an external bias voltage at the drain-source contact which is larger than the built-in potential $U_{DS} > \varphi_{bi}$. The Fermi energy levels in the metals are now separated by eU_{DS}. The band edges of the silicon are slanted upwards from the gold to the titanium and do not feature any parabolic bending. The electric field in the silicon is constant across the whole

core $E = (U_{DS} - \varphi_{bi})/w$. The electron and hole distributions are depicted for each metal separately. The contribution from each carrier species to the total thermionic-emission particle current is affected by the shape of the band edges after the voltage is applied.
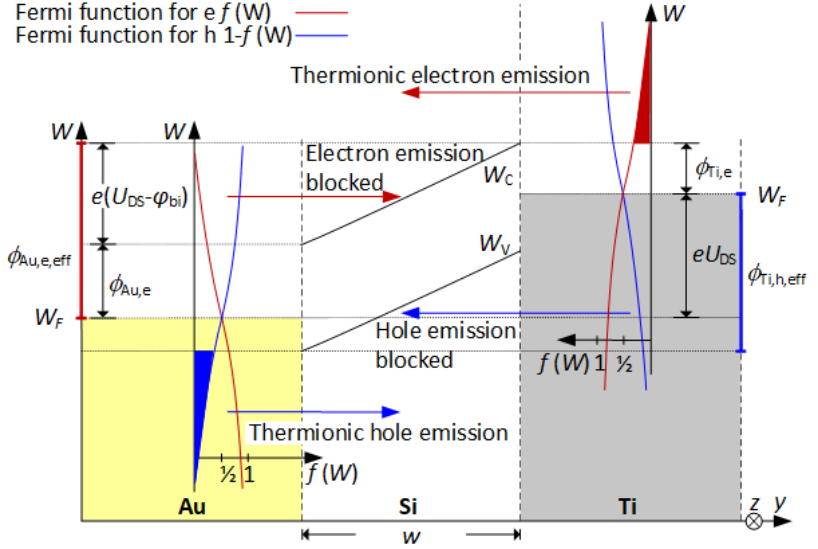
Figure 3. Thermionic emission in a MSM junction at a positive bias voltage U_{DS}. The particle currents for each carrier species in each metal are depicted. The reference potential is on the titanium electrode. The band edges in the silicon are slanted due to the voltage. The tilted band edges affect the ability of carriers in the metals to overcome the barriers and to reach the opposite metal side. Even high-energy electrons in the gold are blocked from reaching the other metal: If they enter the silicon, they are repelled back from the electric field into the gold. The shape of the silicon conduction band edge appears as an additional barrier with an effective barrier height $\varphi_{Au,e,eff}$. Electrons have to feature at least an excess energy of $W_F + \varphi_{Au,e} + e(U_{DS}-\varphi_{bi})$ to reach the titanium. Even then, scattering in the silicon will cause energy loss, further reducing the chances of reaching the titanium. A similar situation applies for the holes in the titanium, which encounter the effective barrier height $\varphi_{Ti,h,eff}$. The holes are impeded from reaching the gold. Conversely, holes in the gold and electrons in the titanium are neither affected by the increased effective barrier height nor by scattering inside the silicon. Neglecting recombination in the silicon, any of those carriers that enters the silicon due to thermionic emission reaches the metal on the opposite side due to the pull of the electric field.

Electrons in the gold with an energy $W > W_F + \phi_{Au,e}$ can enter the silicon. However, the electric field works against their movement and pulls them back towards the gold. The electron emission is blocked by a new effective barrier with a height of approximately

$$\phi_{Au,e,eff} = \phi_{Au,e} + e \times (U_{DS} - \varphi_{bi}). \tag{22}$$

Furthermore, the core width w is larger than the typical mean free path of carriers in the silicon, such that they scatter and lose their energy. Hence, even carriers that could provide sufficient energy to overcome the effective barrier $W > \phi_{Au,e,eff}$ are impeded in reaching the titanium.

Considering the conditions at the entire junction when increasing the applied voltage, electrons in the gold and holes in the titanium do not contribute anymore to the total thermionic emission current. The emission currents of holes in the gold and of electrons in the titanium are not affected.

Photoelectric Emission

Fowler investigated the photoelectric emission from metals into the vacuum and the dependency of the emission on the photon energy $\hbar\omega$ [15]. Electrons in metals face a potential barrier at the boundary between the metal and the vacuum, which prevents them from leaving the metal. The barrier extends until the energy level χ_0, which is the energy level that carriers need to overcome to enter the vacuum: An electron in a metal with an initial energy W absorbs a quantum of light with energy $\hbar\omega$ and is elevated to a higher energy level $W + \hbar\omega$. If the final carrier energy is larger than the total barrier height χ_0, i.e., $W + \hbar\omega > \chi_0$ the carrier can escape the metal and enter the vacuum state. The work function ϕ_W is the energetic separation between the Fermi energy and the vacuum level. It defines the smallest barrier that electrons must overcome to leave the metal. The photon energy must be larger than the work function $\hbar\omega > \phi_W$ to create photoelectric emission, neglecting the spread of the occupation probability from 0.88 to 0.12 within $4kT = 100$ meV around the Fermi energy at room temperature.

Fowler evaluated the number of electrons N_B with sufficient energy to contribute to photoelectric emission for a given metal and photon energy $\hbar\omega$. He considered an electron gas distributed according to the Fermi-Dirac statistics $f(W)$ in the metal. He assumed a square-root dependency for the density of states $N(W)$ similar to Equation (6) and integrated the carrier energy distribution starting from the lowest possible escape energy $W > \chi_0 - \hbar\omega$ up to infinity. He thus found that N_B depends on the work function ϕ_W, the total height of the barrier χ_0 and the photon energy $\hbar\omega$ [15]

$$N_B = \int\limits_{\chi_0 - \hbar\omega}^{\infty} N(W)f(W)\mathrm{d}W \sim \frac{\left(\hbar\omega - \phi_W\right)^2}{\left(\chi_0 - \hbar\omega\right)^{1/2}}. \tag{23}$$

Fowler assumed that the photoelectric emission current I_P under constant illumination is proportional to N_B.

We assume that Fowler's theory for photoelectric emission from metals into vacuum can also be applied to the case of internal photoemission at a metal-semiconductor interface. At the MS interface, the electrons have to overcome the barrier to the semiconductor and enter the conduction band as opposed to the vacuum level. To account for this difference, we exchange the work function by the barrier height in Equation (23), $\phi_W \rightarrow \phi_{B,e}$. For internal photoemission, the total barrier height χ_0 is not a meaningful parameter as it is much larger than the photon energy $\chi_0 \gg \hbar\omega$.

Internal photoemission becomes more complex in an MSM junction, because electrons and holes from either metal may contribute to the overall current, all of which are subject to different barrier heights. We assume that Equation (23) still holds for each MS interface individually, but does not describe the total emission current in the MSM junction. In Appendix B.2, we describe the conditions under which Equation (23) can still be applied to analyze photoelectric currents from Au-Si-Ti junctions.

2.2.3   Carrier Densities in the Silicon Core

We now consider the possible influence of a gate voltage, i.e., the charge carrier densities n and p in the silicon core of the MSM structure in Figure 1 (a) and their dependencies on the applied gate voltage are investigated.

Table 3. Charge carrier densities n and p at the MS interfaces for zero bias, $U_{DS} = 0$ and $U_{GS} = 0$. The effective densities of states are assumed to be $N_C = 3.2 \times 10^{19}$ cm^{-3} and $N_V = 1.8 \times 10^{19}$ cm^{-3}. The temperature is $T = 293$ K and $kT = 25.3$ meV.

Gold-silicon interface	Charge carrier density at the interface
Electron density	$n = N_C \exp\left(-\dfrac{\phi_{Au,e}}{kT}\right) = 2.6 \times 10^5 \, \text{cm}^{-3}$
Hole density	$p = N_V \exp\left(-\dfrac{\phi_{Au,h}}{kT}\right) = 1.9 \times 10^{14} \, \text{cm}^{-3}$
Silicon-titanium interface	Charge carrier density at the interface
Electron density	$n = N_C \exp\left(-\dfrac{\phi_{Ti,e}}{kT}\right) = 7.1 \times 10^8 \, \text{cm}^{-3}$
Hole density	$p = N_V \exp\left(-\dfrac{\phi_{Ti,h}}{kT}\right) = 6.8 \times 10^{10} \, \text{cm}^{-3}$

We begin the analysis for the case without any bias ($U_{DS} = 0$, $U_{GS} = 0$). The silicon core is weakly p-doped with an acceptor concentration of $N_A \approx 10^{15}$ cm^{-3}. Directly at the interfaces, the carrier densities are distinctively different to the densities in the bulk silicon. For gold, the barrier heights $\phi_{Au,e}$ and $\phi_{Au,h}$ pin the separation of the Fermi energy and the edge of the conduction band $W_C - W_F = \phi_{Au,e}$ or the valence band $W_V - W_F = \phi_{Au,h}$, respectively. The same relations hold for titanium using $\phi_{Ti,e}$ and $\phi_{Ti,h}$. Then, the carrier densities at the MS interfaces can be directly inferred from the barrier heights in Table 2 and with the Equations (8) and (9). The resulting densities are shown in Table 3. There is a significant hole carrier density present at the Au-Si interface, which is due to the small separation of the valence band edge and the Fermi energy.

The analysis for the carrier densities in the bulk silicon core under external bias is based on numerical simulations. The details thereof are explained in Appendix C.2. The dependence of the carrier concentration in the silicon core is evaluated in the gate voltage range -10 V $< U_{GS} < +10$ V.
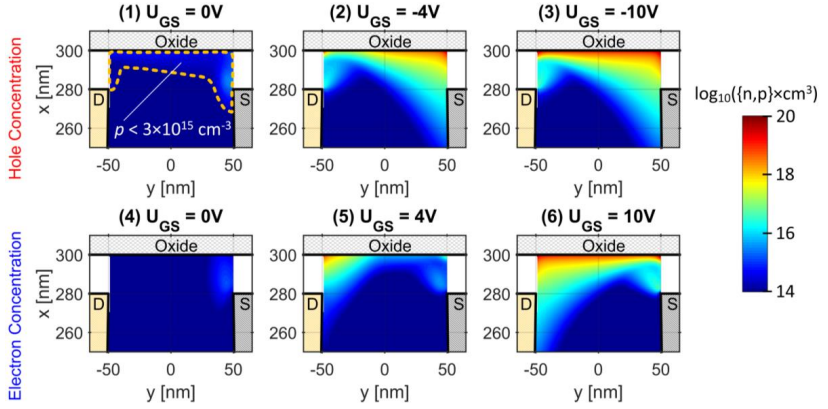
Figure 4. Gate-voltage dependence of the carrier concentration in the silicon core. For all subfigures, the drain-source voltage is $U_{DS} = 1.5$ V. The Subfigures **(1)** through **(3)** depict the hole concentration for the negative voltage range, $U_{GS} = 0$, –4 V, –10 V, while Subfigures **(4)** through **(6)** show the electron concentration for positive voltages $U_{GS} = 0$, 4 V, 10 V. The bulk of the silicon is practically depleted of any mobile carrier concentration for all bias voltages. Subfigure **(1)** shows that even for $U_{GS} = 0$, there is a weak hole accumulation area immediately below the gate oxide, which has been highlighted for better visibility with a dashed orange line. The hole concentration in the highlighted area is between 10^{14} cm^{-3} and 3×10^{15} cm^3. The peak hole concentration in this accumulation layer is two orders of magnitude larger than the concentration in the surrounding bulk semiconductor. The carrier concentration in the channel rises if larger gate voltages are applied so that the channel extends deeper into the silicon, see Subfigures (2) and (3). For electrons (lower row), the slightly p-doped semiconductor must be inverted to form a conductive electron channel. The inversion voltage is around $U_{GS} = 4$ V, see Subfigure (5). At high bias voltages, both the electron and the hole channel sustain a very high carrier concentration of the order of 10^{20} cm^{-3}. At such high voltages, the drain and the source contact are connected by a highly conductive channel.

For the simulation, we assumed a constant drain-source voltage of $U_{DS} = 1.5$ V, while the gate-source voltage is swept. The results of the simulation are shown in Figure 4. The Subfigures (1) through (3), upper row, depict the results for holes and the Subfigures (4) through (6), bottom row, depict the results for electrons. The simulation shows that there is a channel with low p-concentration below the gate oxide already at zero gate voltage, see Figure 4 (1). The conductive channel is established due to a band bending directly at the Si-SiO$_2$ interface, which is present even without any bias. The hole density in the channel, measured at the center between the source and the drain contact ($y = 0$ nm) and at the interface between Si and the oxide ($x = 300$ nm) amounts to $p = 4.5 \times 10^{14}$ cm^{-3}. The channel is constrained to the Si-SiO$_2$ interface. The carrier density drops by one order of magnitude just 20 nm further into the silicon core ($x = 280$ nm). The silicon is otherwise fully depleted of mobile

charge carriers. The simulation results confirm our previously stated assumption of low carrier densities in the bulk silicon.

A non-zero gate bias U_{GS} manipulates the potential $\varphi(x,y)$ in the silicon and controls the carrier densities according to Equations (10) and (11). The subfigures labeled with (1), (2) and (3) in Figure 4 indicate the evolution of the p-channel with a progressively stronger negative gate voltage. The negative gate voltage leads to a hole accumulation at the silicon-gate oxide interface, and the p-channel extends deep into the silicon core. The channel assumes a wedge-like shape due to the bias U_{DS} applied between the drain-source contacts. The hole density at $U_{GS} = -10$ V amounts to $p = 1.6 \times 10^{19}$ cm^{-3} at the center of the channel ($x = 300$ nm and $y = 0$ nm). This constitutes an increase in the charge carrier density of more than four orders of magnitude as compared to the case without a gate voltage. As a consequence, the conductivity $\sigma = ep\mu$ of the channel increases in the same proportion, see Equation (16).

The evolution of the electron concentration for positive gate voltages is indicated in the bottom row of Figure 4, diagrams (4), (5) and (6). At $U_{GS} = 0$ V, the only significant electron concentration is close to the source contact while the electron density is negligible elsewhere. In the range between 0 V$< U_{GS} < +4$ V, the electron concentration increases. Above $U_{GS} \geq 4$ V, an n-channel begins to form, which marks the inversion of the p-doped silicon core. The carrier density and the extension of the n-channel into the silicon increase with higher gate voltage. At $U_{GS} = 10$ V the electron concentration reaches $n = 1.2 \times 10^{19}$ cm^{-3} at the center of the channel. It features the same wedge-like shape as the p-channel at $U_{GS} = 10$ V, but its orientation is flipped.

To summarize, applying a gate voltage creates a conductive channel below the gate oxide in the MSM junction. The type of the channel can be selected by the sign of the gate voltage. The channel increases the conductivity of the MSM junction by several orders of magnitude, depending on the magnitude of the gate voltage. A detailed analysis in the Chapters 5 and D shows that the opto-electronic bandwidth of MSM junctions can be significantly increased if dielectric relaxation comes into play.

2.3   Surface Plasmon Polaritons

Surface Plasmon Polaritons (SPP) are spatial charge density waves that are bound to metal-dielectric interfaces. The plasmonic mode $\vec{\Psi}$ depends on the frequency ω and on the wave vector $\vec{k}_s$,

$$\vec{\Psi}(\vec{r},t) = \vec{\Psi}_0 \exp(\mathrm{j}\omega t - \mathrm{j}\vec{k}_s \cdot \vec{r}). \tag{24}$$

The subscript s is either $s = m$ for the propagation in the metal, or $s = d$ for propagation in the dielectric. Equation (24) can be applied to both, the electric field $\vec{\Psi}_0 = \vec{E}$, or to the magnetic field $\vec{\Psi}_0 = \vec{H}$, respectively. The dielectric function ε_d in the dielectric is assumed to be constant, and the dielectric function in the metal $\varepsilon_m(\omega)$ is frequency-dependent. We model the frequency dependence of the dielectric function in the metal according to a plasma, which consists of free electrons that can move among spatially fixed ions. The concentration of ions is N, and the number of free electrons per ion is f_e. The plasma is assumed to be lossless, so that there are no collisions between the charges as they move. The plasma frequency $\omega_p = N f_e e^2 / (\varepsilon_0 m_e)$ defines the threshold at which charge density disturbances can propagate in the metal. For smaller frequencies $\omega < \omega_p$, there is no propagation and the field is exponentially attenuated. The dielectric function of the metal in this model is (page 12 in [16])

$$\varepsilon_m(\omega) = 1 - \frac{\omega_p^2}{\omega^2}. \tag{25}$$

$\varepsilon_m(\omega)$ is negative for frequencies below $\omega < \omega_p$ and positive otherwise, $\omega \geq \omega_p$.

The inset of Figure 5 shows a plasmonic mode travelling along the z-direction. The z-components of the propagation constant in the dielectric and in the metal are β. This propagation constant β depends on the free-space wave number $k_0 = 2\pi/\lambda_0 = \omega/c$ and on the dielectric functions $\varepsilon_m(\omega)$ of the metal and of the dielectric ε_d (page 26 in [16]),

$$\beta = k_0 \sqrt{\frac{\varepsilon_m(\omega)\varepsilon_d}{\varepsilon_m(\omega) + \varepsilon_d}}. \tag{26}$$

Figure 5 shows the dispersion relation $\omega(\beta)$ of light propagating along the interface of a dielectric and a lossless metal. The frequency range $\omega < \omega_{sp} = \omega_p / \sqrt{\varepsilon_d + 1}$ is the domain of the SPP (blue line in Figure 5). This SPP branch of the dispersion relation is very similar to the light line for small frequencies, where the group velocity is $d\omega/d\beta = v_d$, the speed of light in the dielectric medium. $\omega = \omega_{sp}$ is a horizontal asymptote, where the group velocity approaches 0. There are no propagating modes in the frequency range $\omega_{sp} < \omega < \omega_p$, because β becomes imaginary. For $\omega > \omega_p$, the dielectric function of the metal is positive, and β becomes real again.
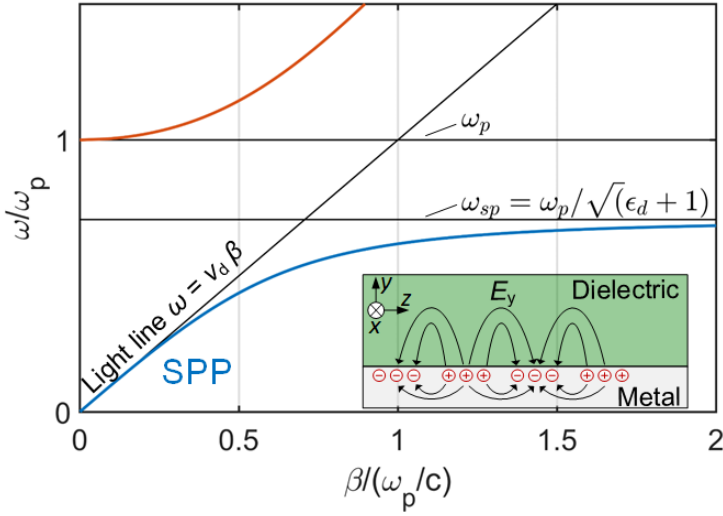
Figure 5. Dispersion relation of light propagating along a metal-dielectric interface. The metal is assumed to be lossless. The domain of the SPP is defined by the frequency range $\omega < \omega_{sp}$. The penetration depth of the electric field in the metal and the dielectric is limited, as shown in the inset. The SPP branch of the dispersion relation is very similar to the light line for small frequencies, and approaches the horizontal asymptote for $\omega = \omega_{sp}$. There are no propagating modes in the frequency range $\omega_{sp} < \omega < \omega_p$. Light is not bound to the interface for frequencies $\omega_p < \omega$.

The transverse propagation constants $k_{m,y}$ (in the metal) and $k_{d,y}$ (in the dielectric) define the electric field decay perpendicular to the direction of propagation for $\omega < \omega_{sp}$ (page 26 in [16])

$$k_{m,y} = \sqrt{\frac{\varepsilon_m^2}{\varepsilon_m + \varepsilon_d}} k_0 , \qquad (27)$$

$$k_{d,y} = \sqrt{\frac{\varepsilon_d^2}{\varepsilon_m + \varepsilon_d}} k_0 . \qquad (28)$$

The quantities $\delta_m = 1/\left|k_{m,y}\right|$ and $\delta_d = 1/\left|k_{d,y}\right|$ are the associated field penetration depths into the metal and the dielectric, respectively.

In the following, we study the instantaneous electric field distribution of an SPP mode travelling along an Au-Si interface in z direction. The chosen free space wavelength is $\lambda_0 = 1550$ nm. We discard the assumption of a lossless metal and assume dielectric functions according to published values. The dielectric functions at $\lambda_0 = 1550$ nm are $\varepsilon_{Au} = -112.7 - j\,6.85$ [17] for gold, and $\varepsilon_{Si} = 11.7$ [18] for silicon, respectively. The metal fills the half-space $y < 0$ nm, and the semiconductor fills the half-space $y > 0$ nm. The interface between the metal and the dielectric aligns with the x-z plane at $y = 0$.

The electric field components of the SPP mode Ψ in x, y and z direction are denominated E_x, E_y and E_z, respectively. Figure 6 schematically depicts the electric field distribution for the field component E_y normal to the interface and it depicts the field component E_z parallel to the interface. The field component E_x is parallel to the interface, but normal to the direction of propagation. It is always zero in this configuration.

The imaginary part of the propagation constant $\text{Im}(\beta)$ determines the attenuation of the SPP due to absorption losses along the axis of propagation. The power attenuation constant α and the associated characteristic absorption length l_{abs} are defined as

$$l_{abs} = \frac{1}{\alpha} = \frac{1}{2\,\text{Im}(\beta)} . \qquad (29)$$
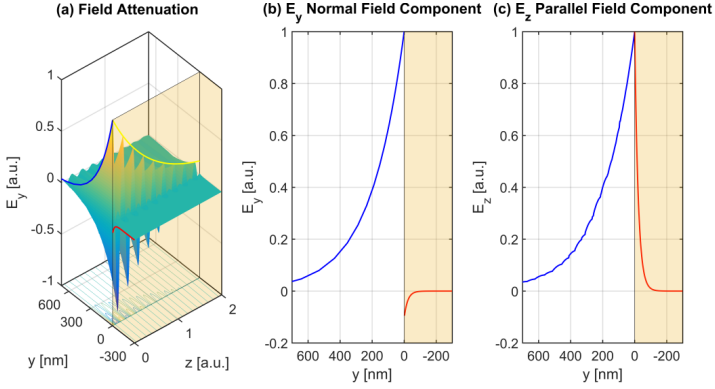
Figure 6. Electric field distribution of the surface plasmon polariton (SPP) for the normal component E_y and the parallel component E_z. **(a)** An SPP with a free-space wavelength of $\lambda_0 = 1550$ nm travels along the z-direction of an MS-interface made of Au and Si. The metal fills the half space $y < 0$ nm, and the semiconductor fills the half space $y > 0$ nm. The metal is highlighted by the golden shading. The field strength in the semiconductor of E_y (normal to the interface) is strongest directly at the interface. The field strength in the metal is significantly less and barely visible in comparison to the field in the semiconductor. The lateral field penetration depths into the materials are very small compared to the wavelength λ_0. The field decays exponentially into the materials (in y-direction), with different penetration depths $\delta_m = 1/|k_m|$ (in the metal) and $\delta_d = 1/|k_d|$ (in the semiconductor). This decay of the normalized magnitude of E_y is drawn in blue for the semiconductor and in red for the metal. The field amplitude along the axis of propagation is exponentially attenuated due to absorption in the metal. The field decay in the direction of propagation is highlighted by the yellow outline. **(b)** Evolution of E_y along a cut line normal to the MS-interface. The electric field changes the sign at the interface due to the imposed continuity of the electric displacement field D_y. Virtually all the field is concentrated in the semiconductor (blue line). The field strength in the metal (red line) is significantly smaller, and the penetration depth is shorter than in the semiconductor. **(c)** Evolution of E_z along a cut line normal to the MS interface. The electric field has the same sign in the metal and the semiconductor, due to continuity of the tangential electric field component.

The field distributions for the normal component E_y in Figure 6 (a), (b), and for the longitudinal component E_z in (c) indicate a strong field concentration at the interface. The lateral extension of the SPP field travelling along an Au-Si interface is given by the combined field penetration depths, resulting in $\delta_m + \delta_d \approx 260$ nm (see Table 4 for details). This is much smaller than the free-space wavelength of λ_0. The SPP is attenuated along the propagation direction, and the power attenuation length is close to $l_{abs} = 13$ µm. The electric field component E_y normal to the interface has a sign change across the interface due

to the continuity of the normal component of the electric displacement field $D_y = \varepsilon E_y$. The field strength in the metal nearly vanishes compared to the field strength in the semiconductor.

Table 4 compares the penetration depths and the power attenuation lengths of a Au-Si interface and a Ti-Si interface, respectively, considering a free-space wavelength of $\lambda_0 = 1550$ nm. A comparison of the two configurations shows that the field penetration into the metal for the Ti-Si interface is much larger and hence the attenuation in the direction of propagation is stronger. Specifically, the characteristic length is only 0.2 µm for the Ti-Si interface, while it is 13 µm for the Au-Si interface. These very short characteristic lengths are a key element of the photodetection concept developed in this work, see Section 4.2.

The large absorption losses of SPP by interaction with the metal are generally seen as a disadvantage of plasmonic devices. However, photodetectors are designed to absorb light as efficiently as possible. The loss property of SPP is ideally suited for a plasmonic photodetector scheme that exploits internal photoemission: The field strength in the metal is the largest directly at the metal-dielectric interface. Hence, the light absorption is the strongest there, too. The light penetration depth into the metal is typically only a few nanometers. This penetration depth is within the reach of the mean free path length of hot carriers. The absorption thus creates hot electrons and holes so close to the interface that the hot carriers may overcome the potential barrier at the interface of the metal and an adjacent semiconductor by tunneling or by thermionic emission.

Table 4. Penetration depths δ_m and δ_d of the SPP electric field, and characteristic length l_{abs} for an Au-Si configuration and a Ti-Si configuration at a free space wavelength of $\lambda_0 = 1550$ nm. The numerical evaluation is based on the Equations (26), (27) and (29) and the dielectric functions for Au [17] und Ti [18].

Interface	Field penetration δ_m	Field penetration δ_d	l_{abs}
Au-Si	24 nm	235 nm	13.3 µm
Ti-Si	84 nm	148 nm	0.2 µm

Furthermore, if the hot carriers do not overcome the barrier directly, they thermalize and create a heated carrier gas, which in turn is able to create a carrier flow by the same mechanisms. The Chapters 4 and 5 discuss how plasmonic structures can be designed to simultaneously achieve record high sensitivities and large opto-electronic bandwidths. Appendix A.2 discusses the light coupling from photonic to plasmonic waveguides and the resulting absorption characteristics.

3 State of the Art in *p-i-n* Photodetectors

Photodetectors are solid-state devices that transfer light signals with optical power P_{opt} to electric currents I_p. They are key elements of optical interconnects, where they represent the interface between the optical and the electronic domain. The operation principle of photodetectors relies on free carrier generation by light absorption at a given optical wavelength λ. The carriers are always generated as electron-hole pairs and increase the conductivity of the absorptive material. The rate of photo-generated electrons equals the rate of the absorbed photons,

$$\frac{I_p}{e} = \eta \frac{P_{opt}}{\hbar \omega}.$$

(30)

$\hbar \omega$ is the energy of a photon at an angular frequency ω. The quantum efficiency η depends on the absorption characteristics of the detector material and the length of the device. Upon generation, the mobile carriers are separated in strong electric fields and drift to the device terminals, where the photocurrent transits to the external circuitry.

The key technical performance indicators for photodetectors are:

1. The opto-electronic sensitivity S (or responsivity R), which defines the efficiency of the optoelectronic conversion. It is given by the ratio of the photocurrent I_P leaving the device and the optical power P_{opt} impinging on the detector,

$$S = \frac{I_p}{P_{opt}} = \frac{\eta e}{\hbar \omega}.$$

(31)

2. The opto-electronic bandwidth is a measure of the maximum useful operating frequency f_{mod}. It indicates to which degree the output photocurrent of the detector can follow a sinusoidally modulated optical power signal with a given modulation frequency f_{mod}. The transfer function $H(f_{mod})$ relates the frequency-dependent photocurrent $I_P(f_{mod})$ to a modulated optical input power $P_{opt}(f_{mod})$,

$$H(f_{mod}) = \frac{I_p(f_{mod})}{P_{opt}(f_{mod})}.$$

(32)

At $f_{mod} = 0$, the transfer function is identical to the DC sensitivity. The 3 dB limiting frequency $f_{3dB, mod}$ indicates where the transfer function has decreased to $\left| H(f_{3dB,mod})/H(0) \right| = 1/\sqrt{2}$ of its initial value at DC. It is used as a figure of merit (FOM) for the bandwidth. The bandwidth is directly linked to the carrier dynamics within the detector. Typical Si/Ge photodetectors reach bandwidths in the range between 10 GHz and 100 GHz [19].

3. The noise equivalent power (NEP) is the optical input power per square-root of the measurement bandwidth $\sqrt{BW}$ that creates an electric signal with a power equivalent to the electronic noise of the device, i.e., the signal-to-noise-ratio (SNR) is 1. The NEP is a measure to which extent the photodetector is useful in handling weak optical signals. The NEP is, e.g., dictated by thermal noise of the resistive elements (Nyquist noise [20]) in the photodetector circuits, which is caused by the statistical movement of carriers, and which may be further increased by amplifiers within the circuit.

A hypothetical ideal photodetector combines high sensitivity and high bandwidth and low noise levels. When considering chip-level optical interconnects, photodetectors should furthermore only require small footprints and fabrication should be suited for seamless integration into CMOS fabrication workflows.

p-i-n Photodetectors

High-speed *p-i-n* photodetectors constitute the state of the art in optical communication technology and serve as a benchmark for comparison with the photodetector concept developed in this work. To this end, the key device characteristics and typical performance parameters shall be discussed in this section.
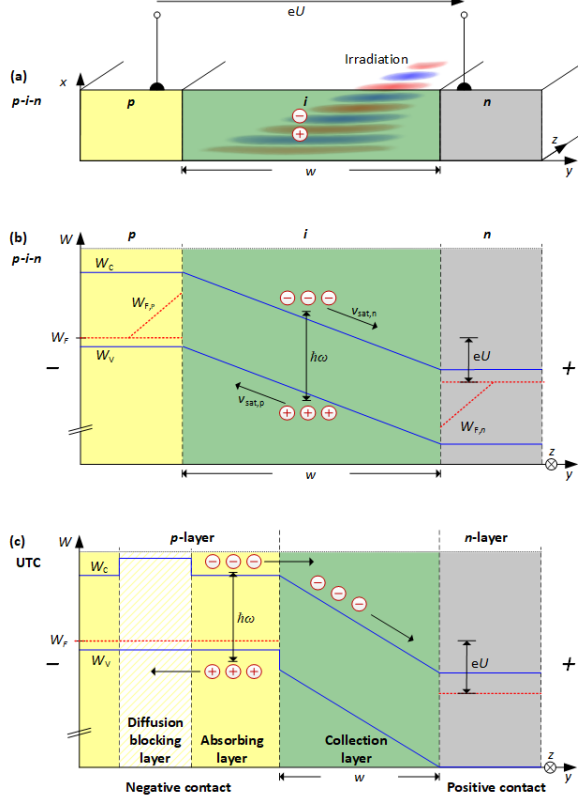
Figure 7. *p-i-n* and UTC photodetector structure and operating principle. **(a)** The intrinsic region of a *p-i-n* detector has a width *w* and is sandwiched between two regions with a high *p*- or *n*-doping. Light impinges onto the device, penetrates into the semiconductor, and creates electron-hole pairs by absorption. **(b)** *p-i-n* band diagram under reverse biasing. A positive voltage *U* is applied to the *n*-layer. The external reverse voltage *U* leads to a potential drop mainly across the intrinsic region. Photogenerated carriers drift to the terminals under the influence of the electric field. The width of the device and the saturation velocity v_{sat} of each carrier type limit the bandwidth of the device. **(c)** Band diagram of a UTC photodiode. Carriers are created in the very thin absorbing layer only. The band gap is too large for absorption in all other layers. Electrons diffuse into the collection layer and drift towards the *n*-layer. Holes cannot enter the collection layer due to an offset in the valence band and diffuse towards the *p*-layer. A diffusion-blocking layer introducing an energy offset in the conduction band prevents electrons from diffusing towards the negative terminal. This avoids long drift distances of the slower holes in the collection layer and leads to very high operating frequencies.

Figure 7 shows a schematic of a typical *p-i-n* detector structure and its operating principle. The detector consists of a *p*-doped ("*p*"), an undoped (intrinsic, "*i*") and an *n*-doped ("*n*") layer, Figure 7 (a). The terminals of the device are attached to the *p*-layer and the *n*-layer. A reverse bias voltage U leads to a potential drop mainly across the intrinsic region and extracts the photocarriers, see Figure 7 (b). To enable the highest dynamic response, the bias voltage U is chosen such that the carriers drift with their saturation velocity v_{sat}. Increasing the voltage further does not increase the bandwidth of the detector. The photocurrent transits through the *p*-doped and *n*-doped regions and exits the device through the terminals.

The undoped intrinsic domain has a typical width from ~ 1 µm (for III-V compound semiconductors) to ~10 – 20 µm (for silicon). This width is chosen according to the light absorption characteristics and the requirements for high-speed operation: If the *p*- and *n*-layers are too close to each other, the quantum efficiency is low, and the bandwidth will be limited by the large capacitance of the junction. If they are too far apart, the drift time of carriers within the *i*-region will limit the bandwidth.

The material-specific absorption coefficient of the semiconductor defines the wavelength band of operation. Table 5 summarizes the performance parameters for various photodetector material systems.

Table 5. Photodetector material systems and typical operating parameters.

Material system / Type	Operating range	Peak sensitivity
Si *p-i-n*	320 nm – 1100 nm	0.7 A/W
Ge *p-i-n*	800 nm – 1800 nm	0.85 A/W
InGaAs *p-i-n*	900 nm – 2550 nm	1.3 A/W
InP/InGaAs UTC	1550 nm	~ 0.3 A/W

A major drawback of *p-i-n* detectors is the bandwidth limitation imposed by the slower carrier species, i.e., the species with the smaller saturation velocity v_{sat}. An optical impulse creates electron-hole pairs in the intrinsic region of width w. The photocurrent is measurable at the device terminals as long as charges flow inside the device. The movement of the slower carriers (typically holes) limits

how fast the device can react to the optical impulse. Furthermore, the concentration of both carrier species reduces the internal field, which is detrimental to the bandwidth.

Uni-travelling-carrier (UTC) photodiodes overcome this limitation by restricting the drift of the slow carrier species. Figure 7 (c) schematically shows the band diagram of a UTC photodiode. Photocarriers are only generated in the absorbing layer, whose width is typically in the order of a few 100 nm. The collection layer is non-absorbing due to the large band gap and serves for extracting the photogenerated carriers.

The UTC design in Figure 7 (c) assumes that electrons are the faster species in the material system. The electron-hole pairs are created within the strongly doped absorbing layer. Due to the strong doping, the electric field is weak and the carriers move by diffusion only. The conduction band edges of the absorbing layer and the collection layer are designed to align well such that photo-created electrons can move easily across the domain boundary. The electrons drift through the collection layer towards the n-layer. The holes cannot enter the collection layer and diffuse only towards the negative contact.

The excelling high-speed capabilities of UTC are due to the specific design of the layer stack: The width of the non-absorbing collection layer is designed such that the p- and n-layers are sufficiently separated, and capacitive limitations of the bandwidth do not become an issue. Furthermore, the collection layer is specifically designed to provide a high electron saturation velocity, which is the main reason for the very high device speed. The absorbing layer is kept short to offer a fast diffusion of holes towards the negative contact at high carrier densities. A diffusion-blocking layer prevents electrons from diffusing towards the negative contact. However, the absorbing layer must still be designed wide enough to ensure a sufficiently high absorption and detector responsivity.

UTC are typically designed for a specific operating frequency. This is necessary to accommodate the constraints imposed by the hollow metal waveguides used for connecting the UTC to external electronics. Typical UTC designs for an operation frequency of 300 GHz reach sensitivities between 0.2 and 0.3 A W^{-1} [21-23].

4 High-Speed Plasmonic Photodetectors

In this chapter, a novel photodetection concept based on internal photoemission is introduced, which we abbreviate with PIPED. The photodetector clearly outperforms similar device concepts, and we show that PIPED are technically viable for chip-to-chip interconnects operating at high speeds and with a high integration density.

We have published this chapter as a paper in *Optica* [J6]. The formatting has been adapted to fit the general format of this thesis. The paper was published with a supplementary information, large parts of which can be found in the Appendix A. Parts of the paper were presented in a talk on the *Conf. on Lasers and Electro-Optics* [C24]. The fabrication steps were presented in a talk on the Conference on *Micro and Nano Engineering* [C28].

The author of this dissertation developed the PIPED device concept together with Manfred Kohl, Christian Koos, Juerg Leuthold and Wolfgang Freude. The author fabricated the devices in the clean rooms of the Institute of Microstructure Technology (IMT) and the Institute of Nanotechnology (INT). The fabrication was supported by Kira Koehnle and Peter Jakobs for the electron-beam lithography. Alban Muslija and the author optimized and applied the dry etching steps. Paul Vincze and the author developed and applied the metallization process at grazing incidence. Yuriy Fedoryshyn supported the fabrication with consultation. The DC-measurements were conducted by the author, Argishti Melikyan and Tobias Harter. The high-frequency measurements and data experiments were additionally supported by Stefan Wolf. The author performed the numerical simulations. The project was supervised by Wolfgang Freude, Juerg Leuthold, Christian Koos and Manfred Kohl. The author and Manfred Kohl, Christian Koos and Wolfgang Freude wrote the paper. All authors revised the paper.

[*start of paper* [J6]]

Silicon-Plasmonic Internal-Photoemission Detector for 40 Gbit/s Data Reception

S. Muehlbrandt[1,2,*], A. Melikyan[1,2], T. Harter[1,2], K. Köhnle[1,2], A. Muslija[1], P. Vincze[3], S. Wolf[2], P. Jakobs[1], Y. Fedoryshyn[4], W. Freude[2], J. Leuthold[4], C. Koos[1,2,*], M. Kohl[1,*]

[1]Institute of Microstructure Technology (IMT), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[2]Institute of Photonics and Quantum Electronics (IPQ), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[3]Institute of Nanotechnology (INT), Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
[4]Institute of Electromagnetic Fields (IEF), ETH Zurich, Zurich, Switzerland
*Corresponding author:

sascha.muehlbrandt@outlook.de; christian.koos@kit.edu; manfred.kohl@kit.edu;

Abstract: Silicon-plasmonics enables the fabrication of active photonic circuits in CMOS technology with unprecedented operation speed and integration density. Regarding applications in chip-level optical interconnects, fast and efficient plasmonic photodetectors with ultra-small footprint are of special interest. A particularly promising approach to silicon-plasmonic photodetection is based on internal photoemission (IPE), which exploits intrinsic absorption in plasmonic waveguides at the metal-dielectric interface. However, while IPE plasmonic photodetectors have already been demonstrated, their performance is still far below that of conventional high-speed photodiodes. In this paper, we demonstrate a novel class of IPE devices with performance parameters comparable to those of state-of-the-art photodiodes while maintaining footprints below 1 µm². The structures are based on asymmetric metal-semiconductor-metal (MSM) waveguides with a width of less than 75 nm. We measure record-high sensitivities of up to 0.12 A/W at a wavelength of 1550 nm. The detectors exhibit opto-electronic bandwidths of at least 40 GHz. We demonstrate reception of on-off keying (OOK) data at rates of 40 Gbit/s.

4.1　Introduction

Chip-level optical interconnects are key to overcome communication bottlenecks in high-performance computing systems [1]. To achieve high data rates and energy-efficient operation, the corresponding transmitters and receivers must be seamlessly co-integrated with CMOS electronics. This is currently impeded by vastly different footprints of electronic and photonic components: Whereas state-of-the-art field-effect transistors can be fabricated [3] with a half pitch of less than 14 nm, typical dimensions of photonic devices amount to tens or even hundreds of micrometers. Plasmonics enables a large reduction of these dimensions by exploiting charge-density oscillations that are strongly confined to metal-dielectric or metal-semiconductor (MS) interfaces [16], thereby leading to short carrier transit times and ultra-small device capacitances. In combination with the large conductivities of metals, any RC limits are hence shifted to very high frequencies, which leads to unprecedented operation speeds [24]. At the transmitter side, ultra-compact plasmonic modulators operating at data rates of 40 Gbit/s and beyond have been demonstrated recently [25, 26]. For the receiver, however, miniaturized plasmonic photodetectors with technically relevant performance still remain to be shown. Previous demonstrations cover photodetectors which exploit passive plasmonic nanostructures for concentrating the incident light to an ultra-small area of an absorbing semiconductor such as germanium [27, 28]. In these approaches, measured sensitivities (responsivities) are typically less than 0.001 A/W due to the limited efficiencies and the intrinsic absorption losses of plasmonic light concentrators. In an alternative approach, the intrinsic absorption of metal can be directly used for photodetection. This is accomplished by internal photoemission (IPE) which exploits photon-assisted transmission of hot carriers across a potential barrier at MS interfaces [15, 29]. Propagation of surface plasmon polaritons (SPP) in plasmonic waveguides is ideally suited for realizing such devices as the plasmonic mode is strongly localized at the MS interface and hence perfectly concentrates the light to the region where absorption leads to the highest generation rate of photo-electrons. The potential of IPE has been demonstrated in several different plasmonic devices, including resonant plasmonic antennas or nano cavities [30-32] and metal-coated silicon waveguides [33-36].

However, the sensitivity of IPE-based photodetectors has so far been limited by the width of the potential barrier that optically excited hot carriers need to overcome. Highest reported sensitivities amount to only 0.019 A/W at a wavelength of 1550 nm [35] – more than one order of magnitude below those of state-of-the-art waveguide-integrated germanium PIN or avalanche photodiodes [19, 37, 38]. Moreover, the bandwidths of these devices are limited to 7 GHz by relatively long drift regions of more than 2 µm.

In this paper, we introduce and experimentally demonstrate a novel approach to exploiting internal photoemission for photodetection. For brevity, we name the resulting device a PIPED, an acronym for plasmonic internal photoemission detector. Our PIPED combines record-high sensitivities of more than 0.12 A/W at 1550 nm with large opto-electronic bandwidths well above 40 GHz. We prove the viability of this concept by receiving on-off keying (OOK) data at a rate of 40 Gbit/s. This is the first demonstration of data transmission at technically relevant bit rates using plasmonic IPE detectors. This superior performance is due to a nano-scale silicon drift region of only 75 nm width, sandwiched between two metal layers with distinctively different light absorption characteristics and Schottky barrier heights. This leads to an ultra-narrow potential barrier and to short transit times under the influence of an external bias voltage. Further simulations and measurements imply that with an optimized coupling approach the sensitivities can be increased at least by a factor of two.

4.2 Operation Principle

The energy band diagram of the Au-Si-Ti photodetector junction is sketched in Figure 8. The junction forms a metal-semiconductor-metal (MSM) light waveguide that guides surface plasmon polaritons (SPP) along the z-direction, which is perpendicularly to the drawing plane of Figure 8. The silicon core has nano-scale dimensions in the order of $w = 100$ nm and constitutes a potential barrier between the metals, thus impeding any charge transfer. The silicon core is weakly p-doped with an acceptor concentration of $n_A \approx 10^{15}$ cm^{-3}. The SPPs dissipate their energy mainly at the Si-Ti interface, because titanium has a larger magnitude of the imaginary part of the complex permittivity $\varepsilon = \varepsilon_1 - j\,\varepsilon_2$ than gold at a wavelength of $\lambda = 1550$ nm ($\varepsilon_{Ti} = 1.6 - j\,30.56$ [18],

$\varepsilon_{Au} = -112.7 - j\,6.85$ [17]). The SPP absorption length along the z-coordinate is below 1 µm (see supplemental Figure 24), therefore the junction length is not truly important and can be chosen in the range $L = (1..20)$ µm, depending on the requirements for the dark current. The absorbed plasmonic wave with angular frequency ω creates hot electrons in the metal within a few nanometers from the interface. The maximum carrier energies exceed the Fermi energy by $\hbar\omega$, Figure 8 (a). These hot electrons have an increased probability of crossing the potential barrier at the metal-semiconductor interface either by thermionic emission or by tunneling, where the tunneling probability depends on the height of the potential barrier and on its width at a given carrier energy [39]. This process is known as internal photoemission [29]. The barrier heights are assumed to be $\Phi_{Au} = 0.82$ eV [12] and $\Phi_{Ti} = 0.62$ eV [14] for the Au-Si and the Si-Ti interface, respectively. This establishes a built-in potential difference across Si core of about $\varphi_{bi} = 0.2$ V in thermal equilibrium (Figure 8 (a)). Without external bias, φ_{bi} impedes electron photoemission from the highly absorbing Ti into the Si layer, and no significant current flow can be measured.

A positive potential exceeding $U = \varphi_{bi}$ is required at the Au electrode to enable photoemission from the Ti layer which is on zero potential, see Figure 8 (b). The silicon layer is much smaller than the depletion layer width of either MS interface due to the weak doping [40]. Hence, the external voltage dominates the voltage drop across the silicon layer for large bias and a constant electric field is present, indicated by a linear increase of the conduction and the valence band energies W_C and W_V along y. The effective barrier width w' at the Si-Ti interface is consequently reduced, and thus the transmission probability of hot electrons is increased. For a silicon core width of $w = 75$ nm and an applied voltage of $U = 3$ V (or $U = 1$ V), the effective barrier width for electrons at the Fermi-level in titanium is reduced to $w' = 17$ nm (or $w' = 58$ nm). The effective barrier width for hot electrons is even smaller and progressively decreases to zero for increasing electron energies, leading to a much larger sensitivity than measured with other IPE-based photodetectors.
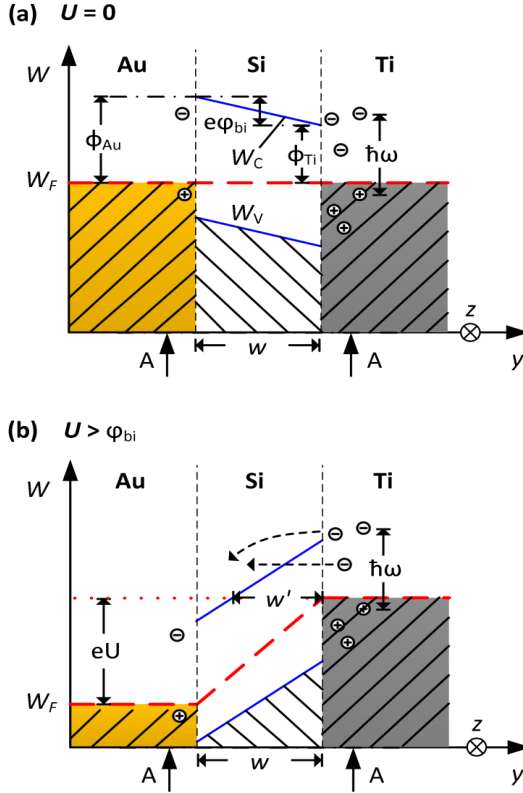
Figure 8. Energy band diagram of an Au-Si-Ti junction. Energy W, lateral direction y as in Figure 9 (b). The silicon layer of width w constitutes a potential barrier impeding charge transfer between the metals. The energy of the Schottky barrier heights are Φ_{Au} and Φ_{Ti}, respectively. W_C denotes the silicon conduction band edge and W_V is the valence band edge. The junction is capable of guiding light in the form of surface plasmon polaritons (SPP), which are mainly absorbed at the Si-Ti interface. Absorbed SPP create hot electrons with a maximum energy of $\hbar\omega$ above the Fermi energy W_F. (**a**) Thermal equilibrium, no bias voltage. The built-in potential φ_{bi} leads to a constant negative electric field along the y-axis inside the silicon layer. Hot carriers created by light absorption at the Si-Ti interface are impeded to cross the barrier by the built-in field. (**b**) Non-equilibrium under applied forward bias voltage U, counted positive in the direction Au → Ti. The voltage drops across the silicon layer. For $U > \varphi_{bi}$ the barrier width w is reduced to an effective width w', thereby increasing the emission probability across the barrier. The section A-A corresponds to associated sections in the device schematics of Figure 9 (b).

Furthermore, the large electric field inside the Si enables high drift velocities. The device structure of the detector is sketched in Figure 9 (a) and (b). Near-infrared light is launched from a silicon photonic strip waveguide into the SPP waveguide junction. The silicon is tapered down from a width of 400 nm in the photonic waveguide to a width of $w = (75 .. 200)$ nm at the MSM junction. The tapered section has a length of 0.55 µm and converts the photonic mode to a plasmonic mode [41]. The silicon core of the detector is wider at its base than at its top as a consequence of the fabrication process. A SiO$_2$ hard mask used for etching the silicon core and a metal top cover are not removed after fabrication. The metal thicknesses on either side of the MSM waveguide are $t = 40$ nm. As an example, Figure 9 (c) and (d) show devices with $w = 200$ nm and a device length of $L = 1$ µm. Thus, the active parts of the device consume less than 1 µm^2 of chip area.

To assess the limitations of the operating speed, we consider several effects: The hot-electron lifetime in metals [42] is only a few 100 fs, which defines the time interval in which transitions from the metal to the semiconductor can occur. The carrier drift time through a 75 nm thick semiconductor layer takes only about 750 fs, assuming a saturation drift velocity [40] of 10^7 cm/s. Similarly, the capacitance of the MSM junction can be estimated based on an equivalent parallel-plate capacitor C across the semiconductor region.

With a metal height of 275 nm, a device length of $L = 5$ µm and a silicon width of $w = 75$ nm, a capacitance below $C = 2$ fF is obtained. This value does not take into account the capacitance of the contact pads. Based on this, an RC time constant of only 100 fs can be estimated by assuming a resistor with $R = 50\ \Omega$ in the bias circuit as depicted in Figure 9 (a). Hence, all these estimates lead to bandwidths of at least several hundred GHz, indicating that compact IPE-based detectors have the potential to overcome existing speed limitations of conventional photodetectors.
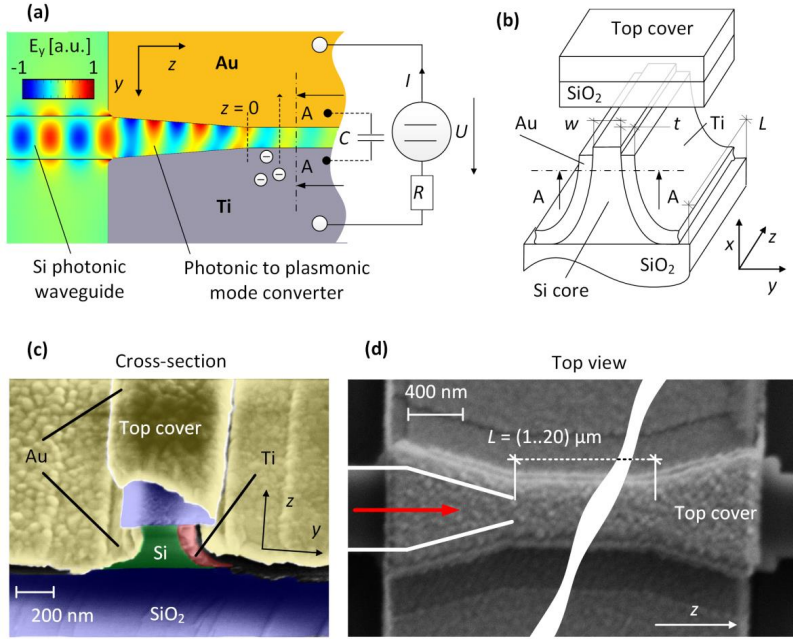
Figure 9. Detector structure and operation principle. **(a)** Light is coupled from a Si photonic waveguide to the Au-Si-Ti junction, which is biased with the external voltage *U*. Absorbed SPPs generate hot electrons which are transferred between the adjacent metals. The section A-A indicates the regions in which the band diagrams in Figure 8 are drawn. **(b)** Schematic of the detector junction with length *L*. The Si core is sandwiched between two metal layers (Au, Ti) of thickness *t*. The silicon core has a height of 300 nm and is wider at the base than at the top as a consequence of the fabrication process. The core width at the top is denoted as *w*. The SiO$_2$ hard mask on top of the Si and its Au cover result from the fabrication process and are only shown in the back half of the structure. **(c)** Cleaved facet of fabricated Au-Si-Ti junction. Due to cleaving, the Au cover on the thermally grown SiO$_2$ hard mask has partially detached. **(d)** Top view of a plasmonic detector. The red arrow denotes the light propagation direction. The Au-Si-Ti junction is hidden below the hard mask and the top metallization.

4.3   Detector Fabrication

The fabrication of asymmetric MSM junctions with a silicon width below 100 nm is challenging, because a short circuit between the electrodes must be avoided. We use electron-beam lithography and isotropic Si etching with SF_6 to create the Si core. Details of the fabrication process are given in the Supplementary Information, see Figure 26. As a first step, a hard mask is structured by using thermally oxidized silicon. By masking and etching, the SiO_2 is removed outside the plasmonic section, leaving a "tabletop" structure above. The silicon sidewalls are selectively metallized by directional evaporation virtually normal to the sidewalls. During metallization, the hard mask prevents a short circuit (see Supplementary Information for details), and silicon widths of 75 nm become feasible. For testing, we fabricate a set of silicon-on-insulator (SOI) chips with different Au-Si-Ti junctions, establishing various PIPEDs with lengths of $L = (1 .. 20)$ µm and silicon core widths of $w = (75 .. 200)$ nm.

4.4   PIPED Characterization

In the following, we report on the results obtained from four different PIPED, which we denote PIPED #1 .. #4. An overview of the physical device properties and the measured performance parameters is given in Table 6. For the sample design and in the fabrication process, we covered a large region of the parameter space, and we optimized the devices for various design goals: The ultra-narrow detector cores of PIPED #1 and #2 were used to demonstrate the high responsivities and the linearity of the device. PIPED #3 is most suitable for polarization contrast measurements, as both polarizations couple with a similar efficiency to the core for this detector width (see Section 4.5). PIPED #4 exhibits the longest carrier drift length and hence was used to demonstrate the high opto-electronic bandwidth, even for the worst-case of a large device. Not all experiments could be done with all devices, as they were driven to their limits and partially damaged during the experiments by applying high bias voltages, or they deteriorated by oxidation of the (not yet protected) titanium. Oxidation can be inhibited by coating the samples with an oxygen-blocking layer [43].

Table 6. Properties of PIPED samples #1..#4 (N. A.: not available)

	Width w	Length L	Sensitivity S(1550 nm)	Bandwidth	Data Exp.
#1	75 nm	5 µm	0.019 A/W (1 V)	N. A.	N. A.
			0.043 A/W (2 V)		
			0.126 A/W (3.25 V)		
#2	75 nm	5 µm	0.042 A/W (1 V)	N. A.	40 GBit/s
			0.102 A/W (2 V)		(1 V)
#3	150 nm	4 µm	0.014 A/W (1 V)	N. A.	N. A.
			0.027 A/W (2 V)		
#4	200 nm	20 µm	N. A.	> 40 GHz	N. A.

The diffraction gratings we used for light coupling excited the photonic waveguide mode with a dominantly horizontal electric field component (see Section 3 in the supplemental material). The total current through the device I comprises a dark current I_d depending only on the bias voltage U, and a photocurrent I_p. The photocurrent increases linearly with the incident optical power P, and the sensitivity $S = I_p/P$ depends on the applied voltage such that $I_p = S(U)\,P$, which is in contrast to conventional p-i-n-photodiodes. Dark current I_d and photocurrent I_p of PIPED #1 (w = 75 nm, L = 5 µm) are characterized at a wavelength of 1550 nm. The total current $I = I_p + I_d$ is measured as a function of laser power P and external DC-bias voltage U (positive polarity from Au to Ti). Figure 10 (a) shows the mean of two subsequently measured I-U characteristics of a detector without and with illumination at a power of P = 310 µW. Both the gold and the titanium electrodes are capable of charge carrier emission into the silicon core. Depending on the sign of the bias voltage U, the resulting current can be positive or negative. As expected, the magnitude of the photocurrent is significantly larger for a positive bias $U > \varphi_{bi}$, which facilitates photoemission across the small potential barrier for electrons generated at the highly absorbing Si-Ti interface.
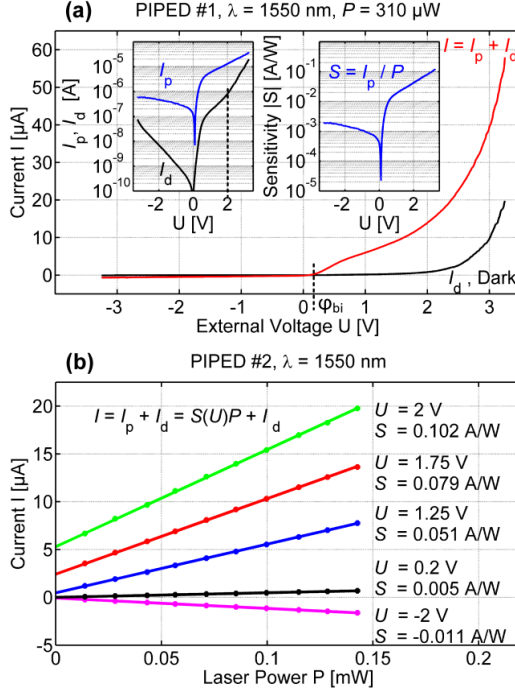
Figure 10. Photodetector DC characterization for PIPED #1 and #2 ($L = 5$ μm and width $w = 75$ nm). **(a)** Total current I for an incident laser power of $P = 310$ μW and dark current I_d as a function of the external (bias) voltage U. For voltages $U > \varphi_{bi}$, the photocurrent is positive, corresponding to carrier injection from the Ti. The photocurrent grows exponentially beyond $U = 1$ V. The insets show a semi logarithmic plot of dark and photocurrent as well as the corresponding sensitivity $S = I_p/P$, exceeding $S > 0.12$ A W^{-1} for a bias voltage of $U = 3.25$ V. The dashed line in the left inset separates the dark current into regions of different exponential growth. As this particular behavior is not present in the photocurrent, we exclude the presence of avalanche multiplication. **(b)** Total device current I vs. laser power P for various bias voltages. The laser power P is measured at the input of the photonic-to-plasmonic mode converter. The filled circles denote the measurements; the solid lines represent linear fits to the measured data.

The left-hand side inset of Figure 10 (a) shows a semi-logarithmic display of photocurrent I_p and dark current I_d, and the inset to the right displays the sensitivity S. For $U > 1$ V, the sensitivity increases exponentially with increasing bias voltage. No saturation is observable. In particular, a positive

photocurrent of $I_p = 38\ \mu A$ is measured for an optical input power of $310\ \mu W$ and a bias voltage of $U = 3.25\ V$. This corresponds to a sensitivity of $S = 0.12\ A/W$, which is more than a factor of 6 higher than the sensitivity of other IPE-based photodetectors [32, 35]. Furthermore, the sensitivity is of the same order of magnitude as values typically measured for comparable state-of-the-art waveguide-based SiGe devices [19, 37, 38]. The steep increase of $S = I_p/P$ for bias voltages $\varphi_{bi} < U < 1\ V$ is caused by the fact that the transmission probability of electrons across the Ti-Si-interface is sensitive to the distinct shape of the potential barrier. A detailed quantitative model is subject of ongoing investigations. Figure 10 (b) shows the total current $I = I_p + I_d = S(U)P + I_d$ as a function of the optical input power for the element PIPED #2 ($w = 75$ nm, $L = 5\ \mu m$). Parameter is the bias voltage. The linear increase of the current with the optical power indicates that the measured photocurrent is due to internal photoemission [15] and not to two-photon absorption in silicon [44]. To examine the influence of carrier multiplication by impact ionization in silicon, we consider the dominant ionization coefficient α_i for electrons. For the maximum bias $U = 3\ V$ (or at $U = 1\ V$) the electric field strength is $E = U/w = 3V/75\,\text{nm} = 40\,V/\mu m$ (or $E = 1V/75\,\text{nm} = 13\,V/\mu m$). The corresponding electron ionization coefficient in silicon amounts to $\alpha_i = 3\mu m^{-1}$ (or $\alpha_i = 10^{-3}\mu m^{-1}$) [40]. This leads to ionization probabilities of $\alpha_i w = 0.23$ (or $\alpha_i w = 0.075$) inside the silicon core. The resulting avalanche multiplication factor $M_0 = \exp(\alpha_i w) = 1.25$ (or $M_0 = 1.08$) is only slightly larger than one, which shows that avalanche multiplication cannot contribute significantly. This finding is supported by comparing the photocurrent I_p to the dark current I_d, see left inset of Figure 10 (a). The logarithmic plot of the photocurrent is essentially a straight line for bias voltages $1\ V < U < 3.25\ V$, while the dark current characteristic has two distinct slopes in the regions $1\ V < U < 2\ V$ and $2\ V < U < 3.25\ V$. This cannot be explained with avalanche multiplication, which acts alike on thermally generated carriers and on photo-generated carriers. Note that the current-voltage characteristics do not follow conventional formulae for MSM-type devices due to the narrow width of the junctions. A detailed quantitative model is subject of ongoing investigations.

4.5 Optical Coupling and Photocurrent Polarization Dependence

The coupling of optical fields to plasmonic MSM-type waveguides is strongly polarization-dependent [16]. Hence, the photocurrent is expected to show a strong contrast between the situations where optical fields with dominantly vertical or dominantly horizontal electric field components are absorbed. Only horizontally polarized optical fields can couple to a pure MSM junction with vertical metallic side walls. To explore the polarization dependence of the photocurrent, we cleaved an SOI chip with PIPED #3 ($w = 150$ nm) and accessed the photonic waveguides from the chip edges using polarization maintaining (PM) fibers. The slow polarization axis of the fiber is aligned with the surface normal of the photonic chip. The orientation of the linear polarization coupled into the PM fiber is then controlled with a half-wave plate, see Supplementary Information for experimental details. The half-wave plate is rotated corresponding to a full 360° revolution of the incoming electric field vector. The linear polarization state at the fiber input is characterized by an angle Θ between the polarization direction and the slow axis of the PM fiber. Figure 11 (a) shows the photocurrent, which is measured as a function of this angle Θ. Note that, due to the strong birefringence of the PM fiber, only the photocurrents at polarization angles of $\Theta = 90° \times$ n, with n = 0, 1, 2, … belong to linear polarization states at the chip input, whereas unknown elliptical polarization states are obtained for any other Θ. Angles of $\Theta_{\mathrm{v}} = 0°$, 180° correspond to a dominant vertical electric field component E_{y}, and angles of $\Theta_{\mathrm{h}} = 90°$, 270° correspond to a dominant horizontal electric field component E_{x}. The photocurrents I_{v} for vertical electric field alignments at angles Θ_{v} are nearly two times as large as the currents I_{h} for horizontal electric field alignments at angles Θ_{h}.

Figure 11. Polarization dependence of photocurrent and simulation of optical fields for PIPED #3 with w = 150 nm. **(a)** Photocurrent I_p as a function of angle Θ between the direction of the linear polarization at the fiber input and the slow axis of the PM fiber with forward bias (blue) and reverse bias (red). The surface normal of the photonic chip is aligned in parallel to the slow axis of the PM fiber. The angles $\Theta_v = 0°$, 180°, 360° correspond to a dominant vertical electric field component E_y, and $\Theta_h = 90°$, 270° correspond to a dominant horizontal electric field component E_x. The insets illustrate the respective electric fields in the photonic waveguide. **(b)** Electric field magnitude at the input of the silicon core, indicated by $z = 0$ in (a), after coupling from a photonic mode with a dominantly horizontal electric field component. The light is localized in the Si core. **(c)** Electric field magnitude at the input of the silicon core ($z = 0$) after coupling from a photonic mode with a dominantly vertical electric field component. Light is localized in the SiO_2 hard mask. Both field distributions deposit the optical power efficiently in the Ti, see Table 7.

50
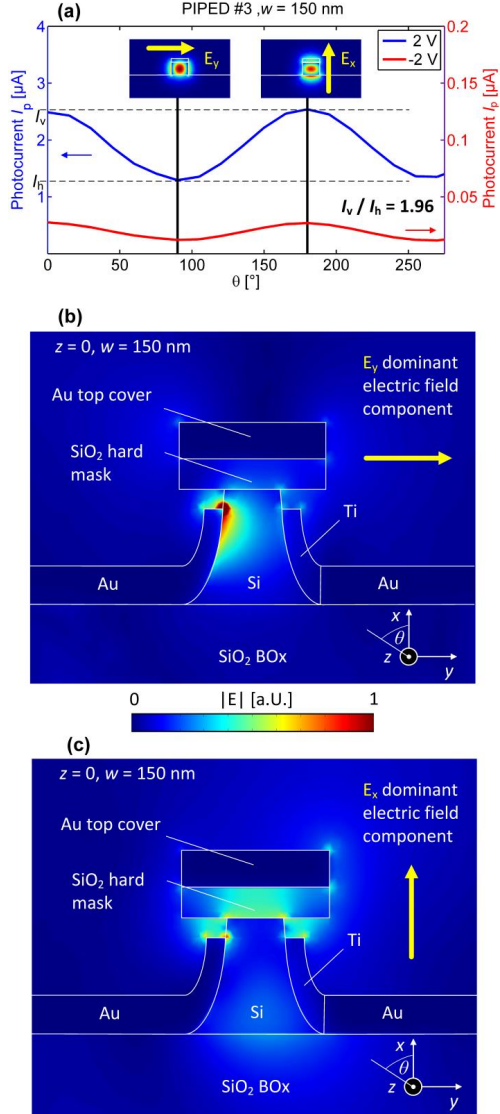
This is surprising at first sight, and needs a detailed investigation of the optical coupling for either polarization. To this end, we use a three-dimensional finite-element method and simulate the coupling of the photonic waveguide modes (dominant E_x or dominant E_y) to the plasmonic photodetector. Figure 11 (b) and (c) show cross-sections of the normalized electric fields in the x-y plane at the input of the silicon core of width $w = 150$ nm, directly at the interface between the core and the tapered mode converter, denoted as $z = 0$ in (a). The optical coupling efficiencies $\eta_{v,h}$ of the dominantly vertically or dominantly horizontally polarized fields are given by referring the power at the input of the silicon core ($z = 0$) to the total power P_0 propagating in the photonic waveguide. These coupling efficiencies $\eta_v = 0.57$ and $\eta_h = 0.50$ are essentially similar for this detector width and hence, the influence of light absorption in the mode converter on the photocurrent polarization contrast can be neglected. However, the field distributions for both polarizations differ considerably. Figure 11 (b) shows that the dominantly horizontally polarized photonic input field mainly couples to the Au-Si-Ti junction and that the field is concentrated inside the silicon core between the two metals. In contrast to that, the dominantly vertically polarized photonic input field cannot couple to the Au-Si-Ti junction as such a mode is not supported by the MSM plasmonic waveguide. However, the junction has a SiO$_2$ hard mask and an Au top cover, which are residuals from the fabrication, see Figure 9 (c) and (d). This layer sequence constitutes an MS interface that supports an SPP mode that couples well to the vertically polarized field of the photonic waveguide. This plasmonic mode confines the electric field to the SiO$_2$ hard mask.

The light absorption in the metals depends on the respective optical field overlap and the imaginary part of the permittivity. The local absorption is in proportion to Im(ε)$\times|E|^2$. While the optical power concentrates on the side of the gold, the field penetration is so small (due to the high conductivity of gold) that absorption as compared to the absorption in titanium is negligible, even if the field strength at the titanium side is small. Table 7 gives an overview of the absorbed power fraction inside the mode converter and inside the detector core for each input polarization. The absorption in the Ti is eight times larger as in the Au for both polarizations, even for a relatively wide silicon layer width of $w = 150$ nm. The vertically polarized plasmonic mode deposits even slightly more optical power in the titanium as compared to the horizontally polarized

mode, for which the structure was designed. In this case, the absorption occurs mostly at the apex of the metallic side walls, where the potential barrier is thinnest and hence the photocurrent is largest.

From our measurements we infer that the PIPED sensitivity can be doubled by choosing the proper polarization. For our best sample PIPED #1 we would hence expect an increase of sensitivity from 0.12 A/W to 0.24 A/W. The currently employed mode converter consumes a significant fraction of the input power but cannot contribute significantly to the photocurrent due to the large separation of the metals compared to the detector core. If the photonic-to-plasmonic converter is redesigned for the specific polarization requirements and for a smaller absorption, even more power would reach the silicon core and hence be absorbed at the Si-Ti interface such that the sensitivity could be further improved.

Table 7. Fraction of absorbed optical power in the photonic-to-plasmonic mode converter (conv.), in the detector core, and total absorbed power per metal contact relative to the total optical input power P_0.

Core width 150 nm	Power absorbed in the Au contact/P_0			Power absorbed in the Ti contact/P_0		
	Conv.	Core	Tot.	Conv.	Core	Tot.
Dominant horizontal polarization	0.073	0.056	**0.13**	0.38	0.42	**0.80**
Dominant vertical polarization	0.018	0.022	**0.04**	0.40	0.43	**0.83**

4.6 Opto-Electronic Bandwidth and Data Reception Experiments

We experimentally quantify the electrical bandwidth of PIPED using a 2-port vector network analyzer (VNA). The VNA stimulus is amplified and drives a LiNbO$_3$ Mach-Zehnder modulator, generating intensity-modulated optical signals up to 65 GHz. A calibrated photodetector (Anritsu MN4765A) with known opto-electronic transfer function is used as a reference. The output pads of the PIPED are contacted with standard RF probes (PicoProbe 40A-GS-100-P). The measured PIPED transfer function (including the RF probe characteristic) is normalized to the value at the lowest modulation frequency of 40 MHz and displayed as a relative sensitivity in Figure 12 (a). Due to bandwidth limitations of modulator and amplifiers, the PIPED current exhibits strong noise for frequencies larger than 40 GHz. Nevertheless, the calibrated transfer function is flat throughout the measured range. Hence, we expect that the RF bandwidth of the device is significantly larger than 40 GHz.

The transfer function exhibits dips at approximately 4 GHz, 7 GHZ and 30 GHz, which are attributed to imperfect connectors of the RF probes. The investigated detector #4 is $L = 20\,\mu$m long and $w = 200$ nm wide. All other devices have smaller dimensions and are expected to show even larger bandwidths. This is confirmed by additional measurements, where we observe comparably flat transfer functions for devices of various lengths and widths. This implies that the devices are neither transit-time limited nor RC-limited.

We expect that the PIPED exhibits this large opto-electronic bandwidth for both input polarizations. The polarization with dominant vertical field alignment creates hot carriers preferentially at the apex of the titanium layer, where the drift distance of the photo-electrons is the shortest. The bandwidth should be slightly higher for this mode in comparison to the mode with horizontal field alignment. However, this has no significant impact on the investigated operating frequency range.
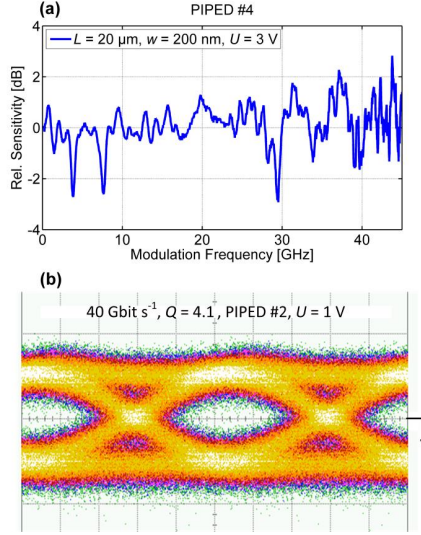
Figure 12. Frequency-dependent photodetector sensitivity and data reception experiment. **(a)** Electro-optic transfer function of sensitivity, normalized to the sensitivity at 40 MHz. The dips at 4 GHz, 7 GHz and 30 GHz originate from reflections at the RF probe. The measurement has been done with the longest and widest device, PIPED #4 **(b)** Eye diagram measured with PIPED #2 for on-off keying at 40 Gbit/s with a measured quality factor of Q = 4.1 and an estimated bit error ratio of BER = 2×10^{-5}. The bias voltage is U = 1 V, the optical power at the input of the detector is P = 1.6 mW. The DC-part of the device current has been removed.

To demonstrate the viability of our photodetector as a receiver of optical data, we perform a data transmission experiment. We generate an on-off keying optical signal with a 40 Gbit/s pseudo-random bit sequence (PRBS) having a pattern length of $2^{31} - 1$. The signal is received with photodetector PIPED #2. We use an available RF amplifier with 50 Ω instead of a transimpedance amplifier, which would be employed in real-world applications. Figure 12 (b) shows the eye diagram at 40 Gbit/s. We use an optical input power of P = 1.6 mW and a bias voltage of U = 1 V at which the dark current is not too large in comparison to the photocurrent so that the electrical signal-to-noise power ratio remains sufficiently large. The DC part has been subtracted. We measure a quality factor of Q = 4.1, corresponding to a bit error ratio [45] (BER)

of 2×10^{-5}. This BER is well below the threshold for a standard second-generation forward-error correction [46] with 7 % overhead.

In summary, we demonstrate the first high-speed and high-sensitivity internal photoemission plasmonic photodetector (PIPED) with a sensitivity exceeding 0.12 A/W at 1550 nm. By changing the state of polarization, the sensitivity can be doubled. We succeeded in data reception of bit rates up to 40 Gbit/s using a PIPED receiver with a core footprint of $5\,\mu m \times 155$ nm, and with a total footprint of less than 1 μm^2. Our experimental findings represent record results for plasmonic photodetectors based on IPE. We believe that compact IPE-based detectors have the potential to overcome existing speed limitations of presently available photodetectors. As such, they could become key components of future high-speed optical transmission systems.

[*end of paper*]

5 Field-Effect Silicon Plasmonic Photo-detector for Coherent T-wave Reception

In this chapter, we expand on the PIPED concept by introducing a further control electrode, which in analogy to a MOSFET is called the gate electrode. Applying a voltage to the gate introduces a field-effect that controls the charge carrier dynamics in the device. We show that optoelectronic bandwidths in excess of 1 THz are feasible using the gate. We have published this research in *Optics Express* [J1]. The formatting has been adapted to fit the general format of this thesis. The paper has been highlighted as an "*Editor's Pick*", which "serve to highlight articles with excellent scientific quality and are representative of the work taking place in a specific field".

The author of this dissertation devised the idea for the field-effect PIPED concept. The photonic and electronic system design was created by the author and Tobias Harter. The author fabricated the devices together with Tobias Harter and Sandeep Ummethala. Andreas Bacher and Lothar Hahn implemented the electron-beam lithography processes. The author conducted the experiments with the support of Tobias Harter, Christoph Füllner, and Stefan Wolf. The author performed numerical simulation of the devices and was supported by Christoph Füllner. The author and Wolfgang Freude analyzed all experimental and simulation data and created the qualitative model of the field-effect PIPED. Christian Koos, Wolfgang Freude and Manfred Kohl supervised the project. The author, Christian Koos and Wolfgang Freude wrote the paper. All authors revised the paper.

[*start of paper* [J1]]

Field-Effect Silicon Plasmonic Photodetector for Coherent T-wave Reception

S. Muehlbrandt[1,2], T. Harter[1,2], C. Füllner[2], S. Ummethala[1,2], S. Wolf[2], A. Bacher[1], L. Hahn[1], M. Kohl[1], W. Freude[2] and C. Koos[1,2,*]

[1] Karlsruhe Institute of Technology, Institute of Microstructure Technology 76344 Eggenstein-Leopoldshafen, Germany

[2] Karlsruhe Institute of Technology, Institute of Photonics and Quantum Electronics, 76131 Karlsruhe, Germany

*Corresponding author: *christian.koos@kit.edu*

Abstract: Plasmonic internal photoemission detectors (PIPED) have recently been shown to combine compact footprint and high bandwidth with monolithic co-integration into silicon photonic circuits, thereby opening an attractive route towards optoelectronic generation and detection of waveforms in the sub-THz and THz frequency range, so-called T-waves. In this paper, we further expand the PIPED concept by introducing a metal-oxide-semiconductor (MOS) interface with an additional gate electrode that allows to control the carrier dynamics in the device and the degree of internal photoemission at the metal-semiconductor interfaces. We experimentally study the behavior of dedicated field-effect (FE-)PIPED test structures and develop a physical understanding of the underlying principles. We find that the THz down-conversion efficiency of FE-PIPED can be significantly increased when applying a gate potential. Building upon the improved understanding of the device physics, we further perform simulations and show that the gate field increases the carrier density in the conductive channel below the gate oxide to the extent that the device dynamics are determined by ultra-fast dielectric relaxation rather than by the carrier transit time. In this regime, the bandwidth can be increased to more than 1 THz. We believe that our experiments open a new path towards understanding the principles of internal photoemission in plasmonic structures, leading to PIPED-based optoelectronic signal processing systems with unprecedented bandwidth and efficiency.

5.1 Introduction

Optoelectronic signal processing techniques offer great potential for generation and detection of waveforms in the sub-THz and THz frequency range, so-called T-waves [6, 47-49], which may be key to ultra-broadband wireless communications in future sixth-generation (6G) networks [7, 8]. In this context, ultra-fast plasmonic photodetectors and electro-optic modulators have been shown to open a path towards generation, detection, and conversion of T-wave signals on the silicon photonic platform [6, 49]. Among the various approaches to plasmonic photodetectors [4, 30, 33, 36, 50, 51], devices based on internal photoemission can combine high bandwidth with a strongly voltage-dependent sensitivity and are thus not only suited for T-wave generation by photomixing, but also for coherent detection of T-waves by optoelectronic down-conversion

with optical local oscillator (LO) signals [6]. However, while the fundamental advantages of plasmonic internal photoemission detectors (PIPED) have been experimentally shown [4, 6], the underlying operating principles are still subject to discussion. One of these questions concerns the relative contributions of photo-emitted electrons and holes to the overall current, which is a key aspect to improve the device performance and to quantify the associated fundamental physical limitations for conversion efficiency and bandwidth.

In this paper, we expand the PIPED concept by introducing a third electrode, which, in analogy to field-effect transistors, we refer to as the gate contact. Varying the gate potential of these field-effect (FE)-PIPED allows to systematically manipulate the contributions of electrons and holes to the overall photocurrent, and thus opens a new experimental access to understanding the device physics. We experimentally demonstrate the concept and verify the underlying principles using dedicated test devices with core widths of 175 nm, and we find that the sensitivity of these devices can be effectively tuned through the gate potential. We further show that the gate voltage increases the sensitivity slope, which determines the down-conversion efficiency for coherent T-wave detection. Specifically, our best performing device shows a two-fold increase in down-conversion efficiency by applying a gate voltage, reaching three times the previously reported efficiency [6]. Based on these findings, we develop a detailed electronic model of the photoemission at the metal-semiconductor interfaces, considering the specific density of states (DOS) on the metal sides. Building upon the improved understanding of the device physics, we further perform simulations showing that the carrier density in a thin channel below the gate can be increased to the extent that the dynamic device response is determined by ultra-fast dielectric relaxation rather than by the carrier transit time. These simulations predict that the bandwidth of the device can be increased beyond 1 THz by applying a gate voltage.

The paper is organized as follows: Section 5.2 describes the device structure and the fundamental operating principles of the FE-PIPED. In Section 5.3, we develop a numerical device model and show how the gate potential influences the carrier density and the device bandwidth. Section 5.4 summarizes the fabrication steps of the devices used in our experiments. In Section 5.5, we describe the measurement of the devices using DC and T-wave signals, and in

Section 5.6 we summarize our findings and formulate design guidelines for PIPED with significantly improved performance for T-wave signal processing.

5.2   FE-PIPED Structure and Operating Principle

The FE-PIPED device structure and the operating principle is shown in Figure 13. Figure 13 (a) and its inset display false-colored scanning electron microscope (SEM) images of the FE-PIPED structure, comprising a plasmonic detector core and a silicon photonic waveguide (WG), which are connected by a photonic-to-plasmonic mode converter. The WG feeds light with optical power P_{opt} to the device. The mode converter transforms the photonic mode in the Si WG to a surface plasmon polariton (SPP) mode inside the detector core. Light is absorbed by the Au drain electrode and by the Ti source electrode, generating hot carriers, which can cross the Si core [4]. This leads to an external photocurrent. A bias voltage U_{DS} is applied via two electrodes adjacent to the Si WG core, which are referred to as drain and source. The drain-source bias increases the efficiency of the crossing of the potential barrier [4]. With a so-called gate electrode on top of the oxide covering the Si core, a second control voltage U_{GS} can be applied between gate and source. We define the source electrode to be at the reference potential, and the drain-source voltage U_{DS} as well as the gate-source voltage U_{GS} are counted positive with respect to the source electrode. The cross-section of the detector core represents a metal-semiconductor-metal (MSM) junction and resembles a MOS-FET phototransistor, which motivates the naming of the electrodes.

Figure 13 (b) shows the schematic cross-section of the PIPED core along with the normalized color-coded electric field profile of the associated SPP mode that is excited by a quasi-TE mode in the silicon feed waveguide. For this type of mode, the dominant electric field component is oriented parallel to the substrate plane. The vacuum wavelength of the incident light is $\lambda = 1550$ nm.
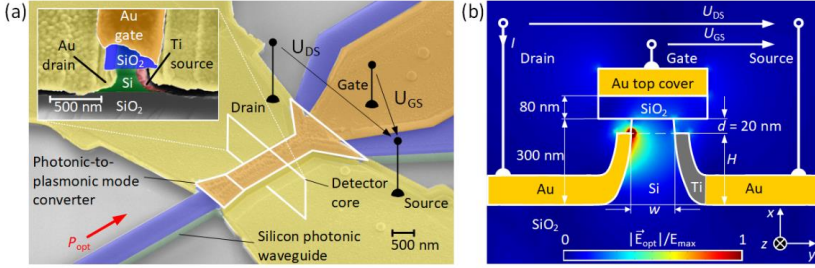
Figure 13. Field-effect PIPED device structure and operating principle. **(a)** False-colored scanning electron microscope (SEM) image of field-effect PIPED. Light with optical input power P_{opt} is coupled to the plasmonic detector core via a silicon photonic strip waveguide (WG) (green) having a thermal oxide on top (blue). A photonic-to-plasmonic mode converter transforms the photonic mode to a surface plasmon polariton (SPP) mode in the core. The inset shows the core cross-section. It consists of a Si core clad with an Au electrode at one sidewall (drain) and a Ti electrode at the other sidewall (source) [4]. The bias voltage U_{DS} is applied via the drain and source electrodes, and the gate voltage U_{GS} is applied via the gate contact (orange). The widening of the gate contact outside the core region facilitates the fabrication process and has no influence on the functionality. **(b)** Schematic cross-section of the photodetector under illumination. The incoming light is coupled to SPP at the interfaces of the Si detector core and the metal layer, and is then absorbed, thereby generating hot carriers. The device core consists of Au and Ti electrodes adjacent to the silicon detector core (metal-semiconductor-metal junction, MSM). The Si core width is chosen to be in the range $w = 75\ldots175$ nm. On top of the silicon core, there is a metal-oxide-semiconductor (MOS) junction. The insulating gate oxide has a thickness of 80 nm and an Au top cover. The source potential is chosen as the reference for the drain-source voltage U_{DS} and for the gate-source voltage U_{GS}. The band diagrams in subsequent figures are taken along a horizontal line indicated by the height H. The vertical separation between the upper edge of the drain and source electrodes and the oxide layer is denoted as d. The magnitude of the computed normalized total optical electric field as excited by a quasi-TE mode in the feeding Si strip WG is color-coded. Note that even though the field in the Si detector core is mainly localized at the Au contact, most of the photons are absorbed in the Ti, which features a larger real part of the complex refractive index than the Au.

The complex refractive indices $\bar{n} = n - \mathrm{j}n_i$ of the Au and Ti electrodes differ significantly both in the real parts, $n_{Au} = 0.32$, $n_{Ti} = 4.0$ [17], and in the (negative) imaginary parts, $n_{i,Au} = 10.6$, $n_{i,Ti} = 3.8$ [18]. Therefore, the SPP mode has an asymmetric shape, and the field tends to localize at the Au-Si interface. Still, the Si-Ti interface dominates the overall absorption in the device, due to the larger real part of the Ti refractive index. The smaller the core width w becomes, the more the absorption length of the lossy SPP mode

decreases. For technically relevant widths of $50\,\text{nm} \leq w \leq 250\,\text{nm}$, the absorption length $1/\alpha$ is below 1 µm, where α denotes the power attenuation coefficient of the SPP mode. A fully vectorial numerical simulation considering the field distribution excited by a quasi-TE mode in a FE-PIPED of core width $w = 100$ nm indicates that six times more power is absorbed in the Ti electrode than in the Au electrode. Accordingly, the bias voltage U_{DS} is chosen such that photo-emitted electrons from the Ti electrode are drawn towards the Au contact, thus forming the main component of the photocurrent in the external circuit. The FE-PIPED can also be operated in a quasi-TM mode where the dominant electric field component is along the normal to the substrate plane. The mode profile and the associated absorption characteristics can be found in the supplementary information of [4].

If a photon with energy hf is absorbed in a metal, it creates an electron-hole pair separated by an energy difference hf, where the actual energy of the carriers depends on their initial energy before the absorption process. As an example, an electron (or hole) which happens to be at the Fermi energy W_{F} will end up after the absorption process as a "hot" electron (or as a "hot" hole) with an electron (or hole) energy $W_{\text{F}} + hf$ (or $W_{\text{F}} - hf$) [15], using the energy counting convention as indicated in Figure 14.

The Schottky barrier at the MS interface influences how high-energy carriers enter the silicon core. If hot carriers have an energy larger than the Schottky barrier height, they overcome the barrier with a probability that is significantly greater than if the carriers have to tunnel through it [39]. The crossing of the potential barrier is known as internal photoemission (IPE) [29].

The energetic distribution of hot carriers before and after absorption depends on the material-dependent density of states (DOS) for Ti and Au [52, 53], respectively, and on the Fermi function. Because at room temperature the Fermi function features a transition from high (0.88) to low (0.12) occupation probabilities within an energetic width of $4kT_0 \approx 100$ meV (Boltzmann constant k, room temperature T_0), the Fermi function can be approximated by a unit step for the photon energies under consideration, $hf = 0.8$ eV.

In the following, we qualitatively explain the influence of the external bias voltages on the band diagram and on the associated photocurrents. A bias

voltage $U_{DS} \neq 0$ leads to a splitting of the quasi Fermi levels for electrons $W_{F,n}$ and for holes $W_{F,p}$ in the silicon core. Independently of the biasing condition, the number of carriers available for photocurrents is generally larger in Ti than in Au as the light absorption is stronger in Ti. The Si band-gap energy amounts to $W_G = 1.11$ eV, and for the barrier height we assume $\Phi_{Au} = 0.82$ eV. Note that this barrier height is much larger than the values of approximately 0.3 eV [54] that are typically found for interfaces of Au and neat Si surfaces. The increased barrier height in our device is due to the fact that the PIPED relies on interfaces between Au and chemically treated Si surfaces prepared by reactive ion etching (RIE). This significantly affects the actual barrier height, for example by gold diffusion or surface states created by the Si processing [13]. Following [12, 13], we estimate a barrier height $\Phi_{Au} = 0.82$ eV, and we experimentally verify this estimate by evaluating the photoemission current [15] of our PIPED devices over a wavelength range between 1270 nm and 1325 nm. To this end, the device was cleaved, and light emitted by a wavelength-tunable laser was coupled directly to the cleaved facet using an optical fiber. By fitting the wavelength-dependent photocurrent to a Fowler's model for photoemission [15], we find a barrier height at the Au-Si interface of $\Phi_{Ti} = 0.82$ eV, which is in good agreement with [12]. We further confirm our result with the physical device simulator SILVACO ATLAS by modeling the PIPED dark current for this barrier height, leading to good agreement with our measurements.

Due to this large Schottky barrier height, photoemission is most efficient for positive drain-source voltages $U_{DS} > 0$, leading to hole emission at the Au-Si interface, Figure 14 (b), (c) and (d). We start with a gate-source voltage $U_{GS} = 0$ which essentially corresponds to a floating gate and therefore to the behavior of a conventional PIPED structure [4, 6]. We then discuss the influence of a changing gate voltage. At the Si-Ti interface, where most of the light is absorbed, the Schottky barrier height amounts to $\Phi_{Ti} = 0.62$ eV [14], leading to electron emission. Because the electron DOS in Ti peaks at $W_F - 0.6$ eV, Figure 14 (a), and because the electron occupation probability below $W_{F,n}$ is close to 1, most of the hot electrons will be found at $W_{F,n}+0.2$ eV, indicated by the dashed black line in Figure 14 (a). Since this energy is still below the barrier energy $W_{B,Ti} = W_{F,n} + \Phi_{Ti}$, internal photoemission predominantly relies on tunneling, and the barrier width at the Si-Ti interface has a significant influence on the electron part of the photocurrent, see Figure 14 (b).
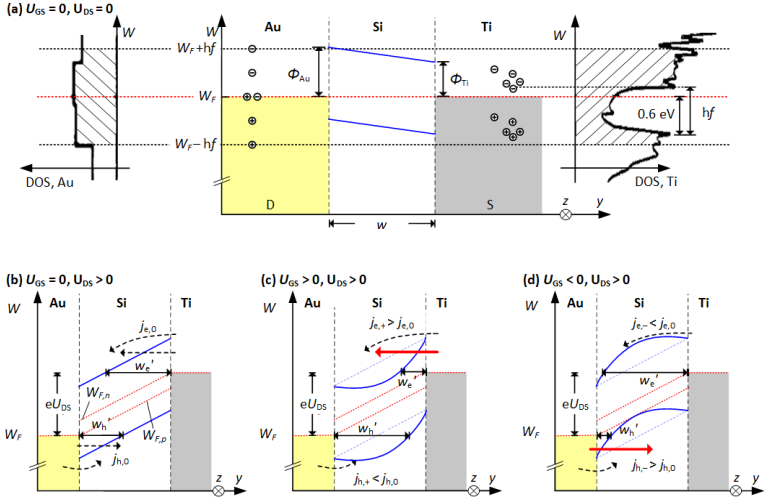
Figure 14. Band diagram of Au-Si-Ti junction for various gate voltages. The Ti electrode (source) is at the reference potential. The band edges (blue) are schematically drawn along a cut line near height H at the upper edge of the metal drain and source electrodes in Figure 13 (b). The difference in the barrier heights Si-Au and Si-Ti leads to a built-in potential $\varphi_{bi} = (\Phi_{Au} - \Phi_{Ti}) / e = 0.2$ V in absence of an external bias. The Fermi energy (red dotted line) is denoted by W_F. Surface plasmon polaritons (SPP) with energy $hf = 0.8$ eV are absorbed in both metals and create electron-hole pairs. **(a)** Thermal equilibrium, $U_{GS} = 0$, $U_{DS} = 0$, with the respective densities of states (DOS) [52, 53]: The electron DOS in Ti peaks at $W_F - 0.6$ eV. Most of the hot electrons will be found at $W_F + 0.2$ eV, below the barrier energy $W_{B,Ti} = W_F + \Phi_{Ti}$. In contrast, the DOS in Au does not strongly depend on energy. Photoemission is therefore governed by hot holes with an essentially uniform energetic distribution. **(b)** Non-equilibrium, $U_{GS} = 0$, $U_{DS} > 0$: Conduction and valence band edges in silicon tilt for a positive voltage U_{DS}, applied between drain (D, Au) and source (S, Ti). Hot electrons from both metals overcome the barriers or tunnel through, leading to particle current densities $j_{h,0}$ (hole emission from Au at $U_{GS} = 0$) and $j_{e,0}$ (electron emission from Ti at $U_{GS} = 0$) and resulting in an external photocurrent. The effective barrier widths w_h' and w_e' limit the carrier emission. The quasi-Fermi levels for electrons $W_{F,n}$ and for holes $W_{F,p}$ are oriented parallel to the band edges. **(c)** Non-equilibrium, $U_{GS} > 0$, $U_{DS} > 0$: Band edges with positive curvature, electrons accumulate under the gate contact. The effective barrier width w_e' at the Si-Ti interface reduces (increased electron emission, $j_{e,+} > j_{e,0}$), while w_h' increases at the Au-Si interface (reduced hole emission, $j_{h,+} < j_{h,0}$). **(d)** Non-equilibrium, $U_{GS} < 0$, $U_{DS} > 0$: Holes accumulate under the gate contact, and the effective barrier width w_e' at the Si-Ti interface increases (electron emission reduced, $j_{e,-} < j_{e,0}$), while the barrier width at the Au-Si interface w_h' reduces (hole emission increases, $j_{h,-} > j_{h,0}$).

Note that from this simplified model illustrated in Figure 14, one may expect that reverting the sign of U_{DS} could lead to a device with similar functionality based on hole rather than electron emission on the Si-Ti interface. However, this effect is not observed experimentally, i.e., the photocurrent for $U_{DS} < 0$ is orders of magnitudes smaller than its counterpart for $U_{DS} > 0$.

In contrast to that, the DOS in Au is rather energy-independent. Photoemission at the Au-Si interface is governed by hot holes that uniformly fill the energy range $(W_{F,p} – 0.8 \text{ eV})...W_{F,p}$. For a Schottky barrier height $\Phi_{Au} = 0.82$ eV and a Si band gap energy of $W_G = 1.11$ eV, the hole barrier is energetically located at $W_{F,p} – 0.29$ eV. For a photon energy of $hf = 0.8$ eV, approximately half of the hot holes are hence to be found at hole energies above $W_{F,p}$, i.e., at electron energies below $W_{F,p}$. As a consequence, a significant fraction of the hot holes which are optically excited in the Au electrode can energetically overcome the barrier irrespective of the exact barrier width or shape. This is distinctively different from internal photoemission of hot electrons from Ti to Si, which largely relies on tunneling, and hence strongly depends on the shape and particularly on the width of the potential barrier. Note that the hot carriers relax and thermalize with the surrounding uniform electron or hole gas if the internal photoemission does not occur within the typical hot-carrier lifetime of a few 10 fs inside the metals [42]. Note also that the contribution of holes from the Au to the overall photocurrent is smaller than that of electrons from the Ti, since the major part of the light is absorbed at the Si-Ti-interface.

In absence of any incident light, the dark current is determined by the emission of cold carriers. The energetic distribution of cold carriers that are in equilibrium with the crystal lattice follows the Fermi-Dirac statistics. Only a small fraction of cold carriers with sufficiently high energy can directly overcome or tunnel through the barriers, a process which is called thermionic emission [39]. Due to tunneling of carriers, thermionic emission also depends on the shape of the respective potential barrier. This leads to a voltage-dependent dark current through the FE-PIPED.

For the band diagram in Figure 14 (b), a positive bias U_{DS} is chosen that overcompensates the built-in potential $\varphi_{bi} = (\Phi_{Au} – \Phi_{Ti}) / e = 0.2$ V. This potential φ_{bi} arises from the different barrier heights for electrons on the gold side (Φ_{Au}) and on the titanium side (Φ_{Ti}). The gate contact is at the same

potential as the source contact, $U_{GS} = 0$ V. In this case, the core width w is much shorter than the depletion region that is generated by the applied drain-source voltage $U_{DS} > 0$ inside the weakly doped Si WG core (acceptor density $n_A \approx 5 \times 10^{15}$ cm^{-3}), such that the space-dependent potential can be described by a linear function. Hence, the applied voltage drops quasi-linearly across the Si core, leading to approximately linear conduction and valence band edges. A non-zero gate-source voltage U_{GS}, Figure 14 (c), (d), influences the potential in the Si WG core only in the immediate vicinity of the gate oxide, similar to the situation for a MOS field-effect transistor. Specifically, an applied gate voltage bends the linear bands such that the effective barrier widths at the Si-metal interfaces are changed, which influences photoemission and hence the device current I. With a fixed drain-source voltage $U_{DS} > \varphi_{bi}$ and for $U_{GS} > 0$, the Si-Ti effective barrier width is reduced to w'_e, and electron injection from the Ti side is increased, Figure 14 (c). Conversely, for $U_{GS} < 0$, the effective Au-Si barrier width is reduced to w'_h, leading to a stronger hole injection from the Au, Figure 14 (d). Moreover, the band bending has a strong impact on the carrier dynamics in the Si core: The electron concentration below the gate oxide is increased for $U_{GS} > 0$, Figure 14 (c), while for $U_{GS} < 0$, the hole concentration increases. If the carrier concentration in the thin channel below the oxide is large enough, signal propagation is dominated by ultra-fast dielectric relaxation rather than by the drift of individual carriers. This leads to a significant increase of the device bandwidth. We explore this feature in more detail in Section 5.3.

The spatial separation d of the gate oxide and the electrodes, as seen in Figure 13 (b), is an important design parameter that strongly influences the device behavior. On the one hand, the distance d must not be chosen to be too large, since this would reduce the influence of U_{GS} on the barrier shapes, and the gate functionality would disappear. Moreover, the photoemitted carriers would be injected far away from the thin conductive channel below the oxide, and the benefits of ultra-fast dielectric relaxation would be lost. On the other hand, choosing d too small would lead to a very small effective barrier width w'_e and would hence have a detrimental impact on the dark current, especially at high gate fields. These considerations lead to a deeper understanding of the measured I-U_{GS} characteristic. We come back to this aspect in Section 5.5.1.

5.3   Optoelectronic Bandwidth

The FE-PIPED has the potential to reach an unprecedentedly high light detection speed. To prove this point, we use a 2D time-domain simulation of an appropriate device model. We use the simulation tool SILVACO ATLAS for the FE-PIPED model Figure 13 (b) with core width of w =100 nm. ATLAS is a physical device simulator that relies on the drift-diffusion model and solves the continuity equations. The influence of the metal-semiconductor interfaces on the device currents are taken into account. Our simulation domain is defined in the x-y plane, i. e., in a cut-plane whose normal vector is perpendicular to the direction of the photocurrent and parallel to the direction of light propagation. With this setting, we simulate the potential $\varphi(x,y)$ and the spatial carrier concentration for various biasing conditions.

Figure 15 (a) shows the electron and hole concentration in the silicon core, directly at the interface of gate oxide and silicon, in the middle between the drain and the source contact where the carrier concentration due to the gate potential is highest. For a drain-source voltage of U_{DS} = +1.5 V and a gate-source voltage of U_{GS} = 0 V, without an applied gate, the hole concentration is in the order of 10^{15} cm^{-3}. When applying a gate voltage as small as U_{GS} = –5 V, holes accumulate, and their concentration increases to 5×10^{18} cm^{-3}. For creating a similar electron accumulation under the gate oxide, the silicon conduction type must be inverted. The inversion threshold is near U_{GS} = 3 V. The electron concentration reaches 10^{19} cm^{-3} for U_{GS} = 10 V. We explore the optoelectronic bandwidth as a function of the gate voltage. To this end, we define a time and space dependent electron-hole pair generation rate $g(t,x,y)$ in the silicon next to the apex of the Au electrode (drain) or, alternatively, next to the Ti source contact. We leave the generation rate $g(t,x,y)$ independent of U_{GS} and thus exclude the influence of the gate voltage on the tunneling and therefore on the injection rate. With forward bias U_{DS} = 1.5 V, the electron-hole pairs created close to the Au contact separate. The electrons will immediately vanish into the metal, and the remaining holes move under the influence of the electric field and create a current between the device terminals.
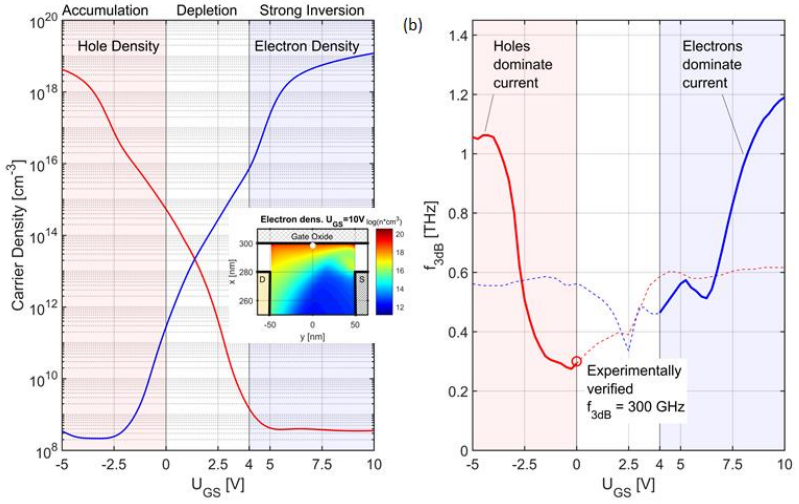
Figure 15. Simulated carrier densities and optoelectronic bandwidth of FE-PIPED at a drain-source voltage $U_{DS} = 1.5$ V. **(a)** The carrier densities are displayed at the interface between the silicon and the gate oxide, in the middle between the source and the drain contact (white circle in the inset). The hole density (red curve) increases with negative gate voltages. The silicon is strongly inverted for gate voltages $U_{GS} > 4$ V, and a significant electron density is created (blue curve) in a channel below the oxide. When increasing the modulus of U_{GS}, the electron and hole densities increase up to 10^{19} cm^{-3}. The inset shows the shape of the electron channel for a strong bias $U_{GS} = 10$ V, which is asymmetric due to the drain-source bias voltage. **(b)** 3 dB limiting frequency for the case of a dominating hole current (solid red curve, $U_{GS} < 0$ V), and for a dominating electron current (solid blue curve, $U_{GS} > 4$ V). The bandwidth of the FE-PIPED is predicted to increase beyond 1 THz at a gate voltage of $U_{GS} < -4$ V, and up to 1.2 THz at $U_{GS} > 10$ V. The dashed curves indicate the respective 3 dB limiting frequencies when hot minority carriers propagate through the majority-type channel. The red circle at $U_{GS} = 0$ V marks the measured 3 dB frequency of our PIPED in a previous publication [6], where the device has the same structure as in our present simulation. Simulation and measurement differ in this point by only 0.3 %. However, there are kinks in the curves which stem from the fact that the transfer functions cannot be accurately approximated by simple RC lowpasses with well-defined limiting frequencies, which is especially true for frequencies below 0.5 THz.

This simulates the dynamic behavior of a light-induced hot-hole injection from the Au electrode into the silicon. An electron-hole pair created at the Ti contact approximates a light-induced hot-electron injection. The creation of an electron-hole pair instead of a single charge is a constraint imposed by the ATLAS simulation environment.

The generation rate g is defined as a Gaussian function both in time and space. The temporal and spatial widths are $2\sigma_t = 100\ \text{fs}$ and $2\sigma_{x,y} = 5\ \text{nm}$, respectively. The spatial extension is small compared to the core width w, and the temporal impulse provides sufficiently high frequency components well above 10 THz. The magnitude of the generation rate g is adjusted such that the resulting current in a 1 µm long device corresponds to typically measured PIPED current in the µA region. The current response to the generation impulse can be Fourier transformed and results in a transfer function $H(f, U_{GS}, U_{DS})$, relating the frequency-dependent photocurrent $I_P(f)$ to a (hypothetical) modulated optical input power $P(f)$,

$$H(f, U_{GS}, U_{DS}) = \frac{I_P(f, U_{GS}, U_{DS})}{P(f)}. \tag{33}$$

We then extract the 3 dB limiting frequency, which we name (optoelectronic) bandwidth. Figure 15 (b) shows the dependency of the FE-PIPED bandwidth on the gate voltage. The solid red curve indicates the bandwidth for a hole current which dominates for $U_{GS} < 0$ according to Section 5.2. The solid blue curve stands for the bandwidth if an electron current dominates. This happens after strong inversion and when the band bending favors hot electron tunnel injection from the Ti for $U_{GS} > 4\ \text{V}$. Our measurements confirm that this condition is actually met, Section 5.5.1, Figure 17 (b).

Our simulations predict a PIPED bandwidth of 290 GHz for gate-source voltage $U_{GS} = 0$, red circle in Figure 15 (b). This is in good agreement with the measured bandwidth evaluated from a similar device described in a previous publication [6], where we found a bandwidth of 300 GHz. Applying a gate voltage increases the bandwidth gradually, because the bandwidth is determined first by the carrier transit time, but is gradually superseded for larger carrier concentrations by the much faster dielectric relaxation time. For large gate voltages $U_{GS} > 10\ \text{V}$

(which is tolerated without dielectric breakdown) the bandwidth increases to 1.2 THz.

Note that the reported increase in bandwidth from 0.3 THz to 1.2 THz of the FE-PIPED could in principle be also achieved by reducing the core width of a standard (without gate and transit-time limited) PIPED from $w = 100$ nm to 25 nm. However, the fabrication of a PIPED with such small dimension would be impractical: The optical coupling efficiency to such a small plasmonic waveguide would be very low, so that the sensitivity decreases significantly. Furthermore, the fabrication would become extremely challenging.

Our results indicate that the FE-PIPED has indeed the potential to outperform conventional Si-Ge photodetectors if T-waves are involved.

5.4 Device Layout and Fabrication

We fabricate and characterize FE-PIPED with gate contacts. The devices are prepared using a similar technique as described previously [4]. We repeat the essential fabrication steps for the convenience of the reader. We start with a silicon-on-insulator (SOI) substrate with 340 nm-thick device layer on top of a 2 µm-thick buried oxide (BOX). As a first step, the gate dielectric is created by thermally oxidizing the top surface of the silicon device layer. The thickness of the oxide amounts to 80 nm, consuming 40 nm and leaving 300 nm of the Si device layer. This approach ensures excellent homogeneity of the oxide layer along with high resilience to dielectric breakdown. In future devices, the oxide could be replaced by a high-κ dielectric layer for further reducing the gate voltage. After oxidation, we coat the substrate with a negative-tone resist and structure the oxide layer with electron-beam lithography and dry etching to obtain a hard mask to fabricate the silicon optical waveguides. We use a CHF_3 etch, which is highly selective and consumes the oxide only, leaving the silicon unharmed. The second step comprises an undercut of the oxide hard mask in the region of the FE-PIPED device core and the adjacent mode converters using isotropic reactive ion etching (RIE) technology. In this step, we use an SF_6 etch at room temperature, which provides a high selectivity of etching silicon and leaving the oxide unharmed. This undercut creates a silicon strip with a wider oxide cap, see inset of Figure 13 (a). In the third step, we use a PMMA mask

and angled evaporation to cover the sidewalls of the silicon strip with Au or Ti. During this step, the oxide cap in combination with the evaporation angle defines the spatial separation d of the gate oxide and the electrodes, see Figure 13 (b).

By design, the PIPED provides two independent optical inputs from opposite sides, but only one of them is required for a photodetector. The remaining one is hence abandoned and covered with a larger gold strip, which connects the gold cap with a remote contact gate pad, see Figure 13 (a) and Figure 16 (b). This electrode is created by Au evaporation normal to the substrate plane, which finally completes the metallization steps. A final anisotropic etch step with SF_6 at -110°C defines the optical waveguides.

Note that the devices used for the experimental exploration of underlying operating principles were purposefully designed to have rather large core widths of $w = 175$ nm, thereby allowing for a clear distinction between electron and hole currents. This, however, leads to a reduced sensitivity as compared to previously demonstrated PIPED with core widths of $w = 75$ nm [4]. The decrease of sensitivity for wider cores can be understood by considering the contributions of the two distinct metal-semiconductor interfaces to the overall photocurrent: For an essentially fixed penetration depth of the optical field in the metal electrodes, the proportion of power propagating in the metals becomes the larger, the smaller the core width w is. However, the Au electrode essentially guides the light, while the Ti electrode absorbs the light. If the electrodes are far apart, the light guidance effect at the Au-Si interface dominates and is essentially not affected by the Ti. If the electrodes are narrowly spaced, absorption is prevalent because the photons guided by the Au-Si interface are drawn towards the strongly absorbing Ti electrode. Specifically, for a core width of 100 nm, the Ti absorbs six times more power than the Au. In addition, the barrier height at the Ti-Si interface is smaller than its Au-Si counterpart, thereby leading to more efficient electron emission at the Ti electrode. At the same time, the reduction of the effective barrier width by applying a drain-source voltage is the larger, the smaller the core width is, i.e., the impact of the drain-source voltage on the carrier emission efficiency is stronger for smaller w. Hence, small core widths w lead to higher sensitivity.

5.5 Experimental Setup and Results

To experimentally verify the expected device behavior, we simultaneously measure the DC and the T-wave characteristics of the PIPED using the setup shown in Figure 16 (a). Two continuous-wave optical tones at frequencies f_1 and f_2 emitted by a pair of tunable distributed-feedback lasers (DFB1, DFB2) are superimposed using a 3 dB coupler and produce a power beat $P_{opt}(t)$ at the difference frequency $f_{THz} = |f_1 - f_2| = \omega_{THz}/(2\pi)$,

$$P_{opt}(t) = P_0 + P_1 \cos(\omega_{THz}t). \tag{34}$$

At the transmitter, this power beat feeds a reverse-biased InGaAs high-speed *p-i-n*-photomixer (Tx-*p-i-n*), the photocurrent of which drives a transmitter antenna. At the receiver, the T-wave is captured by a second antenna and coupled to the drain-source contact of the FE-PIPED, where it is superimposed with a static drain-source bias voltage U_{DS} using T-wave chokes, see Figure 16 (b). This leads to a modulation of the voltage-dependent FE-PIPED sensitivity *S,* the amplitude of which, within a small-signal approximation, can be assumed to be proportional to the amplitude $\hat{U}_{THz}$ of the received T-wave signal.

The FE-PIPED is fed with the same power beat as the *p-i-n*-photomixer. The photocurrent is then given by the product of the time-dependent sensitivity and the time-dependent incident optical power, both of which oscillate at the same frequency f_{THz}, thus leading to a homodyne down-conversion of the incoming T-wave signal using the optical power beat as a local oscillator, see Section 5.5.1 for a more detailed description.
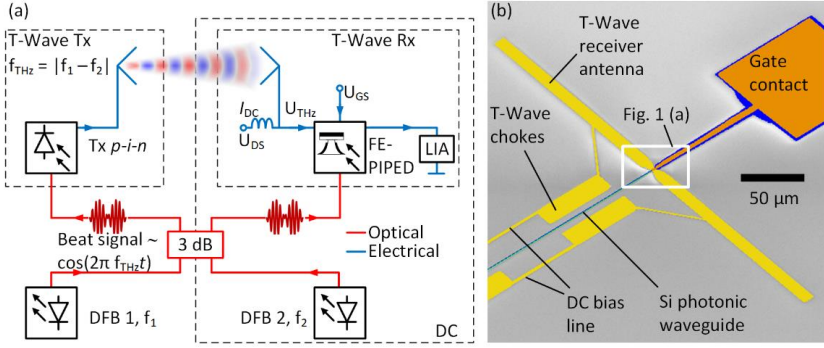
Figure 16. Characterization scheme for FE-PIPED. **(a)** Two DFB lasers emitting near a vacuum wavelength of $\lambda_0 = 1535$ nm are tuned to slightly different optical frequencies f_1 and f_2. These emissions are superimposed to generate a power beat at the desired THz frequency, $f_{THz} = |f_1 - f_2|$. This power beat leads to a photocurrent $i(t) \sim P(t)$ in the Tx p-i-n-photodiode, which drives an antenna emitting the resulting T-wave (T-Wave Tx). At the T-wave Rx, the T-wave is coupled to the drain-source contact of the FE-PIPED, which is additionally biased by a DC voltage U_{DS} via on-chip T-wave chokes. The incoming T-wave voltage leads to a time-dependent modulation of the FE-PIPED sensitivity S with a modulation amplitude that is essentially proportional to the amplitude $\hat{U}_{THz}$ of the T-wave signal. The FE-PIPED is fed with the same power beat that is used to generate the T-wave signal, and the PIPED-photocurrent is coupled to a transimpedance pre-amplifier (TIA) via a bias-T. The TIA is not drawn for simplicity. This photocurrent partly results from the product of the time-dependent sensitivity and the time-dependent incident optical power, both of which oscillate at the same frequency f_{THz}, thereby enabling homodyne down-conversion of the T-wave signal. A lock-in amplifier (LIA) in combination with an amplitude modulation of the THz signal is used to distinguish the down-converted signal from the other components of the photocurrent. The experiment is repeated at various biasing conditions for the drain and the source of the FE-PIPED. **(b)** Scanning-electron microscope image of the PIPED T-wave receiver. The receiver antenna couples the incoming T-waves to the PIPED, which is optically fed via silicon photonic WG. Electrical contacts (DC bias line) are used to supply the DC bias voltages. T-wave chokes decouple the T-waves from the DC bias.

To distinguish the down-converted signal from the other components of the FE-PIPED photocurrent, we use a lock-in technique. To this end, we superimpose an amplitude modulation onto the T-wave signal by varying the reverse bias at the Tx p-i-n photodiode at a frequency of 12 kHz. A lock-in-amplifier (LIA) is then used to extract the corresponding signal component from the output of the transimpedance amplifier (TIA). The FE-PIPED is electrically connected with standard DC probes (GGB Picoprobe MCW-23-2820-1) to a two-channel

precision source measure unit (SMU, Keysight B2900A), which supplies the drain-source voltage U_{DS} and the gate-source voltage U_{GS}. In addition, the SMU records the DC part I_{DC} of the FE-PIPED current I through a bias-T. Parts of this setup (transmitter, antennas, and LIA) belong to a commercially available THz spectroscopy system (Toptica Photonics, TeraScan 1550). For measuring the static FE-PIPED characteristic, we employ only the components enclosed in the dashed rectangle marked "DC" in Figure 16 (a). The optical power beat is coupled to the active part of the FE-PIPED through a grating coupler (GC), a silicon photonic waveguide (WG), and a photonic-to-plasmonic mode converter, see Figure 13 (a) and Figure 16 (b). The T-wave signal is so weak that it has no influence on the DC part of the photocurrent. The SMU measures this average DC produced by the optical power beat.

5.5.1   DC Characterization of FE-PIPED

The total FE-PIPED output current $I(U_{DS}(t),U_{GS})$ consists of the dark current $I_D(U_{DS}(t),U_{GS})$ and the photocurrent $I_P(P_{opt}(t),U_{DS}(t),U_{GS})$. The sensitivity $S(U_{DS}(t),U_{GS})$ depends on the applied voltages and describes the dependence of the photocurrent on the optical power (in the following we drop the time-argument for better readability),

$$I_P\left(P_{opt},U_{DS},U_{GS}\right)=S\left(U_{DS},U_{GS}\right)P_{opt}. \qquad (35)$$

The total average current can be written as

$$I_{DC}=\langle I_P\rangle+I_D=S\langle P_{opt}\rangle+I_D, \qquad (36)$$

where $\langle x(t)\rangle$ denotes a time average of the harmonic quantity $x(t)$ over one period of the T-wave oscillation. The DC sensitivity S is measured by recording the average current with and without illumination, and by relating the difference to the optical input power,

$$S=\frac{\langle I_P\rangle}{\langle P_{opt}\rangle}=\frac{I_{DC}-I_D}{\langle P_{opt}\rangle}. \qquad (37)$$

Figure 17 shows the DC sensitivity of a FE-PIPED having a core width of $w = 175$ nm as illustrated in Figure 14, with drain-source and gate-source voltages U_{DS} and U_{GS} as measurement parameters. Figure 17 (a) visualizes the data as a color-coded plot with equidistant contour lines. For a drain-source voltage difference smaller than the built-in potential, $U_{DS} \leq \varphi_{bi} = 0.2$ V, hot carriers entering the silicon core are dragged back to the respective metal electrodes, and the sensitivity is close to zero. For bias voltages $U_{DS} > \varphi_{bi}$, carriers emitted into the silicon core are pulled towards the opposite metal electrode, and the photocurrent increases with U_{DS}. Two distinct areas with a large sensitivity can be identified. In the upper left region of Figure 17 (a) with $U_{GS} < 0$ V and $U_{DS} > 1$ V, hole emission from the gold electrode is favored by the band bending, which decreases the width of the Schottky barrier as depicted in Figure 14 (d). The upper right area with $U_{GS} > 6$ V and $U_{DS} > 1$ V is associated with strong electron emission from the titanium electrode as illustrated in Figure 14 (c). These areas are separated by a trench near $U_{GS} = 4$ V where the sensitivity is small. Figure 17 (b) depicts the sensitivity along horizontal cut lines in Figure 17 (a), the color-coding corresponds to the coding of the horizontal lines in Figure 17 (a).

The dependence of the sensitivity on the gate voltage is not symmetric and reflects the interplay of different barrier heights and widths for the Au-Si (hole injection) and the Si-Ti interfaces (electron injection), of the influence of the different DOS, and of the photon absorption differences of titanium and gold, see Section 5.2 and Figure 14 (a). This is the reason why the minimum sensitivity is not found for $U_{GS} = 0$ V. In summary, we see that the sensitivity is greatly increased by a gate voltage $U_{GS} \neq 0$ V. The largest sensitivities are to be seen for $U_{GS} > 0$ V due to the stronger impact of a gate voltage on the electron emission at the Ti electrode. The maximum measured sensitivity of 60 mA/W is obtained for $U_{DS} = 1.5$ V and $U_{GS} = 9$ V, and it exceeds the value for $U_{GS} = 0$ V by more than a factor of four. The minimum sensitivity is found near $U_{GS} = 4$ V. Note that the FE-PIPED discussed here was intentionally designed to have a large width $w = 175$ nm of the Si core, see Section 5.4 for details. The wide core also reduces the dark current and allows for a wider range of gate voltages, which was instrumental to analyzing the full characteristics of the device. This, however, comes at the price of a reduced photoelectric sensitivity of only 15 mA/W without gate voltage, and 60 mA/W with an applied gate

voltage. Both values are smaller than the sensitivity 120 mA/W previously obtained for a PIPED without a gate contact [4], but with a core width of only $w = 75$ nm.

We also tested FE-PIPED with a gate contact and smaller core widths, leading to sensitivities in the same range as reported before [4]. In particular, we have characterized an FE-PIPED with a core width of $w = 75$ nm, reaching a sensitivity of $S = 120$ mA/W with an applied gate voltage of $U_{GS} = -3$ V and at a drain-source voltage $U_{DS} = 2.5$ V. In our previous publication [4] for a similar device without a gate, we had to apply a higher drain-source voltage of $U_{DS} = 3.25$ V to reach the same performance. We find that a gate voltage does not significantly increase the sensitivity for narrow-core devices.
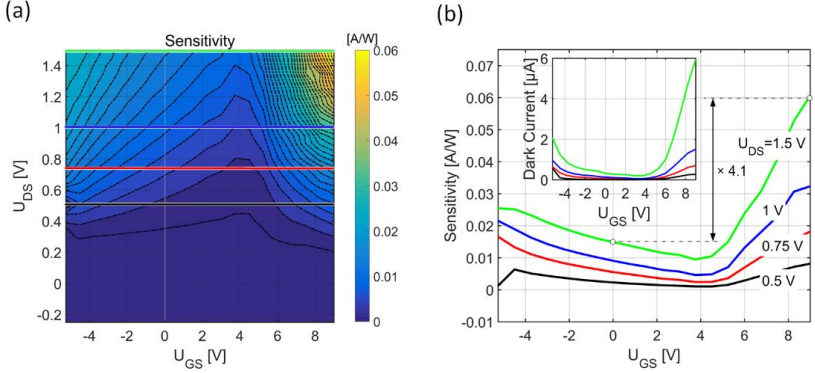
Figure 17. Measured DC sensitivity for a FE-PIPED with core width $w = 175$ nm, operated at an average optical power of $P = 0.21$ mW in the silicon photonic feed WG. **(a)** Color-coded contour plot of sensitivity S as a function of drain-source voltage U_{DS} and gate-source voltage U_{GS}. For $U_{DS} < 0.5$ V, no significant sensitivity can be measured for any gate voltage. For $U_{DS} > 0.5$ V, two distinct areas of large sensitivity can be identified. The left area corresponds to hole emission from the Au electrode for $U_{GS} < 0$, Figure 14 (d), whereas the right area corresponds to electron emission from the Ti electrode for $U_{GS} > 0$, Figure 14 (c). In between, near a gate voltage of $U_{GS} = 4$ V, the photocurrent is minimum. **(b)** FE-PIPED sensitivity as a function of U_{GS} for constant drain-source bias voltages U_{DS} of 0.5 V, 0.75 V, 1 V, and 1.5 V, corresponding to the horizontal lines in (a). The maximum measured sensitivity of 60 mA/W is obtained for $U_{DS} = 1.5$ V and $U_{GS} = 9$ V and exceeds the value for $U_{GS} = 0$ V by more than a factor of four ($\times 4.1$). The minimum sensitivity is found near $U_{GS} = 4$ V. The inset displays the dark current with the same color coding as the main graph.

We attribute this result to the fact that the narrow core allows for an efficient tunneling injection already at moderate drain-source biases, and the further band bending due to the gate voltage does not significantly improve the tunneling efficiency as is the case for larger core widths. Note, however, that the gate voltage can improve the sensitivity slope $\mathrm{d}S/\mathrm{d}U_\mathrm{DS}$ of the FE-PIPED, which is key for an efficient T-wave down-conversion. This aspect is discussed in more detail in Section 5.5.2.

The measured dark current of the FE-PIPED is illustrated in the inset of Figure 17 (b). It follows the same general behavior as the photocurrent, with a minimum at $U_\mathrm{GS} = 4$ V, followed by a steep rise for large negative or positive gate voltages. For $U_\mathrm{DS} = 1.5$ V, the dark current is 320 nA without an applied gate voltage and it increases under the influence of the gate voltage to 2 µA at $U_\mathrm{GS} = -5$ V and to 6 µA at $U_\mathrm{GS} = 9$ V. While both the photocurrent and the dark current show a rise for larger gate voltages, the photocurrent increases significantly stronger.

For a noise estimate, we concentrate on $U_\mathrm{GS} = 9$ V with a dark current $I_\mathrm{D} = 6$ µA and a sensitivity $S = 60$ mA/W. To obtain a lower limit for the noise-equivalent power (NEP), we assume shot noise of the dark current to be the dominant noise source. Note the restrictions implied in our assumption: The two metal-semiconductor junctions are both subject to a random transition of carriers, which may be correlated – we neither account for this effect nor for any resulting frequency-dependence of the noise power spectral density. With these assumptions, we find a current power spectral density for the shot noise of the dark current of $2\mathrm{e}\,I_\mathrm{D} = 1.9 \times 10^{-24}\,\mathrm{A}^2/\mathrm{Hz}$, leading to an $\mathrm{NEP}_\mathrm{shot,\,dark} = \sqrt{2\mathrm{e}\,I_\mathrm{D}}/S = 23\,\mathrm{pW}/\sqrt{\mathrm{Hz}}$. We compare this result with the noise of an external impedance-matching resistor $R = 50\ \Omega$, having a noise current power spectral density of $4\mathrm{k}T_0/R = 3.2 \times 10^{-22}\,\mathrm{A}^2/\mathrm{Hz}$ at room temperature $T_0 = 290\,\mathrm{K}$, which leads to an $\mathrm{NEP}_\mathrm{Nyquist} = \sqrt{4\mathrm{k}T_0/R}/S = 300\,\mathrm{pW}/\sqrt{\mathrm{Hz}}$. Because $\mathrm{NEP}_{50\Omega} \gg \mathrm{NEP}_\mathrm{shot}$, we conclude that the dark current of the FE-PIPED does not significantly contribute to the overall noise as soon as the device is embedded into a 50 Ω-terminated T-wave circuit.

5.5.2 T-wave Characterization of FE-PIPED

If the drain-source voltage $U_{DS}(t)$ is modulated by a T-wave signal, the sensitivity $S(U_{DS}(t), U_{GS})$ will be time-dependent as well. Exploiting the fact that the photocurrent according to Eq. (35) is given by the product of the time-dependent sensitivity and the optical power, we may then use a time-dependent power beat $P_{opt}(t)$ according to Eq. (34) as a local oscillator (LO) for homodyne down-conversion of the T-wave signal to the baseband. In our experiment, the drain-source voltage consists of a strong DC bias $U_{DS,0}$ and a weak sinusoidal T-wave voltage $U_{THz}(t)$, which originates from the receiver antenna and which is superimposed onto the DC bias $U_{DS,0}$ by T-wave chokes, see Figure 16 (b). The drain-source voltage then reads

$$U_{DS}(t) = U_{DS,0} + U_{THz}(t), \qquad U_{THz}(t) = U_1 \cos(\omega_{THz}t + \varphi). \quad (38)$$
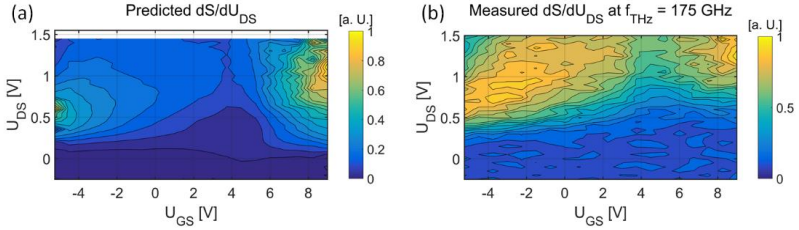
Figure 18. Color-coded normalized derivative $dS/dU_{DS} \propto I_{BB,I}$ as a function of U_{DS} und U_{GS}. **(a)** Evaluation of the DC data from Figure 17 (a). Two plateaus with a normalized value close to 1 are to be seen, and a trench with a normalized value close to 0 for $U_{GS} = 4$ V. **(b)** Measured normalized values from recording the baseband current amplitude $\hat{I}_{BB}$. The general similarity of (a) and (b) is obvious. The influence of the field-effect on the sensitivity is maintained even at T-wave frequencies.

For a sufficiently small T-wave amplitude U_1, we may linearize the function $S(U_{DS}(t), U_{GS})$ in a small-signal approximation in the vicinity of the operating point $U_{DS,0}$,

$$S\left(U_{DS}(t), U_{GS}\right) = S_0\left(U_{DS,0}, U_{GS}\right) + \left.\frac{dS}{dU_{DS}}\right|_{U_{GS}} U_{THz}(t). \tag{39}$$

By combining Eq. (34), (35), (38) and (39), the baseband part of the photocurrent can be written as

$$I_{BB} = I_{BB,0} + I_{BB,1}, \qquad I_{BB,1} = \frac{1}{2}\left.\frac{dS}{dU_{DS}}\right|_{U_{GS}} U_1 P_1 \cos\left(\varphi(\omega_{THz})\right). \tag{40}$$

The phase $\varphi(\omega_{THz})$ describes the phase difference between the received T-wave, Eq. (38), and the phase of the power beat Eq. (34), and is adjusted for maximum $I_{BB,1}$ by slightly varying the THz frequency. Note that in Eq. (40) the direct current $I_{BB,0}$ can be separated from the mixing product $I_{BB,1} \propto \hat{P}_1(t)$ by modulating the transmitted THz power and using a LIA, see Figure 16 (a) and [6].

According to Eq. (40), the detected baseband current amplitude $I_{BB,1}$ is in proportion to the sensitivity slope dS/dU_{DS} for a given gate-source voltage U_{GS}. Evaluating the measured DC characteristic $S_0(U_{DS,0}, U_{GS})$ of Figure 17 (a), we compute this derivative and display the result in Figure 18 (a). Two peaks of dS/dU_{DS} with a normalized value close to one can be seen, and we find a trench with a dS/dU_{DS} close to zero for $U_{GS} = 4$ V. Applying a gate voltage hence improves the efficiency of T-wave down-conversion and thus increases the mixing product $I_{BB,1}$. This is experimentally confirmed by directly measuring the baseband current amplitude $I_{BB,1} \propto dS/dU_{DS}$ as a function of U_{GS} and U_{DS}, see Figure 18 (b). The similarity to Figure 18 (a) is evident: There is a distinct maximum at $U_{GS} > 6$ V and $U_{DS} > 1$ V, a trench at $U_{GS} = 4$ V, and another plateau-like maximum towards smaller values of U_{GS}. Note that, for both of these maxima, the measured improvement of the down-conversion efficiency relative to the case $U_{GS} = 0$ V is not as pronounced as predicted by the evaluation of the DC characteristics. This aspect is subject to further investigations.
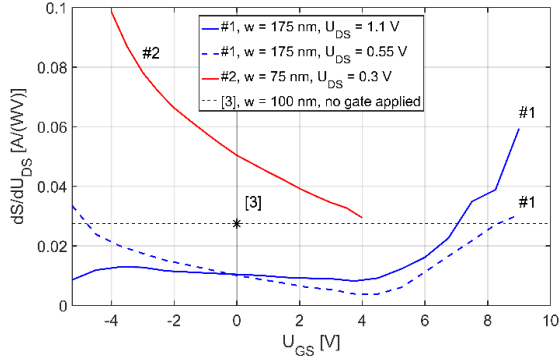
Figure 19. Gate-source-voltage dependent sensitivity slope $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}} \propto I_{\mathrm{BB,1}}$ as measured for FE-PIPED with widths w of 75 nm and 175 nm. The sensitivity slope is evaluated from the measured DC characteristics $S(U_{\mathrm{DS}}, U_{\mathrm{GS}})$ according to Eq. (39). The solid and dashed curves denoted with #1 belong to the same device which also forms the basis of the analysis in Figure 17 and Figure 18. The two curves represent horizontal cut lines through the two pronounced maxima of Figure 17 (a) at $U_{\mathrm{DS}} = 0.55$ V (dashed) and $U_{\mathrm{DS}} = 1.1$ V (solid). The slope is largest for bias voltages of $U_{\mathrm{DS}} = 1.1$ V and $U_{\mathrm{GS}} = 9$ V and reaches $S_1 = 59$ mA/(WV). This value is six times larger than the one obtained for the same drain-source voltage and $U_{\mathrm{GS}} = 0$ V. It is also more than two times larger than the standard PIPED design used in [6], which is indicated with *. This device featured a width of 100 nm, but lacked a gate electrode and was hence operated without any gate-source voltage. The curve labeled #2 belongs to another FE-PIPED with a core width of $w = 75$ nm. For this device, the maximum sensitivity slope $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}} = 98\,\mathrm{mA}/(\mathrm{W\,V})$ at $U_{\mathrm{GS}} = -4$ V is two times larger than the sensitivity slope at $U_{\mathrm{GS}} = 0$ V and three times larger when compared to the standard PIPED design used in [6].

Figure 19 compares the gate-voltage-dependent sensitivity slope $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}}$ for FE-PIPED with core widths in the range 75 nm $\leq w \leq$ 175 nm. Without any gate voltage, $U_{\mathrm{GS}} = 0$ V, the sensitivity slope $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}}$ becomes larger with decreasing width w. An applied gate-source voltage can strongly increase $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}}$ for each width considered. The sensitivity slope increases S with decreasing core width w. However, the dark current also grows with U_{DS} and could even destroy the device when a DC voltage is applied. This affects especially the devices with a small core width w. Hence, the explored range of gate-source voltages was limited. The largest $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}}$ in our experiments is measured for a width of $w = 75$ nm and reaches $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}} = 98$ mA/$(\mathrm{W\,V})$ for bias voltages of $U_{\mathrm{DS}} = 0.3$ V and $U_{\mathrm{GS}} = -4$ V. This is two times the sensitivity

81

slope without a gate voltage and a value more than three times larger when compared to our previously highest reported sensitivity slope $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}} = 27 \ \mathrm{mA}/(\mathrm{W\,V})$ for a device without gate and a core width of $w = 100$ nm [6]. Note that the device #2 in Figure 19 was only measured up to a gate-source voltage of 4 V so that the increase of the sensitivity towards larger positive gate voltages is not visible. We believe that a systematic optimization of mode converter, core width, oxide thickness and choice of metals will allow to further increase the sensitivity slope and thus the efficiency of optoelectronic down-conversion of T-wave signals in FE-PIPED.

5.6   Design Considerations

Based on our investigations, we believe that the performance of PIPED can be greatly improved both in bandwidth and in sensitivity by the addition of a gate contact. The key design parameters of the resulting FE-PIPED are the width w of the Si waveguide core and the vertical separation d between the upper edge of the drain and source electrodes and the oxide layer. Specifically, the photodetector sensitivity is the larger, the smaller the core width w becomes. We find that FE-PIPED with a width of $w = 75$ nm provide the highest sensitivity slope $\mathrm{d}S/\mathrm{d}U_{\mathrm{DS}}$. However, in this case also the dark current increases so that a compromise has to be found. The feasibility of FE-PIPED with even smaller widths $w < 75$ nm is under investigation.

For the spatial separation d between the gate dielectric and the upper edge of the metal drain and source contacts an optimum value has to be found that avoids a practical short while still maintaining the effectiveness of the gate.

The T-wave down-conversion efficiency is significantly increased by a positive (large w) or a negative gate voltage (small w), which then leads to a more efficient hole injection for $U_{\mathrm{GS}} < 0$ on the Au side, or to a more efficient injection of electrons for $U_{\mathrm{GS}} > 0$ on the Ti side.

So far, the T-wave input circuit to the FE-PIPED feeds the T-wave to the drain-source contact leading to a sensitivity slope > 0.1 A/(WV). One could also think of coupling the T-wave signal to the gate-source contact. This, however, would reduce the sensitivity slope to a much lower value. The magnitude of the sensitivity slope for the case of a modulated gate-source voltage can be derived

from the slopes of the family of curves in Figure 17 (b). For a device with a width of $w = 175$ nm, we expect a sensitivity slope below 0.01 A/(WV) throughout the complete range of gate-source bias voltage and for the maximum drain-source bias of $U_{DS} = 1.5$ V. However, this finding is constrained to FE-PIPED with a thermal oxide cap, as used in our experiments. The usage of a high-κ dielectric layer between the silicon and the gate contact might lead to a more favorable potential distribution in the silicon, thereby leading to a stronger effect of the gate voltage.

We expect that future work in these directions leads to improved systems for generation and detection of sub-THz and THz waveforms on the silicon photonic platform.

5.7 Conclusion

We introduce an extended concept for a field-effect-(FE-)controlled plasmonic internal photoemission detector (FE-PIPED). The field-effect opens a new experimental access to the understanding of internal photoemission in plasmonic structures and of the associated carrier dynamics. Based on an improved understanding of the device physics, we perform simulations and show that the gate voltage can enhance the bandwidth of FE-PIPED to beyond 1 THz. We show experimentally that the gate field also increases the sensitivity slope, which is a key figure of merit for T-wave coherent reception.

[*end of paper*]

6 Silicon–Plasmonic Integrated Circuits for THz Signal Generation and Coherent Detection

In this chapter, we show that PIPED can operate as both, T-wave transmitters and receivers, and hence lend themselves for application in T-wave systems. We suggest a proof-of-concept design, in which we monolithically integrate PIPED with T-wave circuits. The research in this chapter has been published in *Nature Photonics* [J4]. The formatting has been adapted to fit the general format of this thesis. The original publication was supplemented by a methods section and a dedicated supplementary information, both of which can be retrieved online https://www.nature.com/articles/s41566-018-0237-x. They have not been carried over in this thesis to limit the length.

The author of this dissertation and Tobias Harter, Wolfgang Freude and Christian Koos developed the idea. The author, Tobias Harter, Sandeep Ummethala and Lothar Hahn contributed to the fabrication of the devices. The author, Tobias Harter, Wolfgang Freude and Christian Koos developed the mathematical formulation. The author, Tobias Harter and Alexander Schmid conducted the measurements. Tobias Harter and Simon Nellen performed experiments to calibrate the reference Tx and Rx. Tobias Harter performed electromagnetic simulations. Wolfgang Freude and Christian Koos supervised the project. Tobias Harter, Wolfgang Freude and Christian Koos wrote the paper. All authors revised the paper.

[*start of paper* [J4]]

Silicon-plasmonic integrated circuits for terahertz signal generation and coherent detection

T. Harter[1,2], **S. Muehlbrandt**[1,2], S. Ummethala[1,2], A. Schmid[1],
S. Nellen[3], L. Hahn[2], W. Freude[1], and C. Koos[1,2]

[1] Karlsruhe Institute of Technology, Institute of Photonics and Quantum
    Electronics, 76131 Karlsruhe, Germany
[2] Karlsruhe Institute of Technology, Institute of Microstructure Technology
    76344 Eggenstein-Leopoldshafen, Germany
[3] Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute,
    10587 Berlin, Germany

Abstract: Optoelectronic signal processing offers great potential for generation and detection of ultra-broadband waveforms in the terahertz range (so-called T-waves). However, fabrication of the underlying devices still relies on complex processes using dedicated III–V semiconductor substrates. This severely restricts the application potential of current T-wave transmitters and receivers and impedes co-integration of these devices with advanced photonic signal processing circuits. Here, we demonstrate that these limitations can be overcome by plasmonic internal-photoemission detectors (PIPEDs). PIPEDs can be realized on the silicon photonic platform, which allows exploiting the enormous opportunities of the associated device portfolio. In our experiments, we demonstrate both T-wave signal generation and coherent detection at frequencies up to 1 THz. To prove the viability of our concept, we monolithically integrate PIPED transmitters and receivers on a common silicon chip and use them to measure the complex transfer impedance of an integrated T-wave device.

6.1 Introduction

Terahertz signals (T-waves) offer promising perspectives for a wide variety of applications, including high-speed communications [47, 55, 56], microwave photonics [57], spectroscopy [9, 58], life sciences [59, 60], as well as industrial metrology [61, 62]. Optoelectronic signal processing techniques are particularly attractive for both T-wave generation [47, 63, 64] and detection [48, 65], especially when broadband tunability of the terahertz frequency is required. On a conceptual level, optoelectronic generation of continuous-wave (c.w.) terahertz signals relies on mixing two optical signals oscillating at frequencies f_a and f_b in a high-speed photodetector, for which the photocurrent depends on the incident optical power [63]. The photocurrent oscillates with a difference frequency $f_{THz} = |f_a - f_b|$ in the terahertz region, which can be relatively easily adjusted over the full bandwidth of the photodetector by frequency-tuning one of the two lasers. In many practical applications, the optical signal oscillating at f_a carries an amplitude or phase modulation, whereas the optical signal at f_b is simply a c.w. carrier. In this case, the phase and amplitude modulation of the optical carrier is directly transferred to the T-wave carrier. This concept shows great potential for high-speed wireless communications at terahertz carrier frequencies and has been at the heart of a series of transmission experiments, in which record-high data rates of 100 Gbit/s and above have been reached [55, 66, 67]. Similarly, optoelectronic techniques can be used for detection of T-wave signals. In this case, the T-wave signal is applied to a high-speed photoconductor and the optical power oscillation at the difference frequency $f_{THz} = |f_a - f_b|$ is used as a local oscillator (LO) for coherent down-conversion to the baseband [48, 65]. This technique was initially developed for frequency-domain terahertz spectroscopy systems offering a widely tunable frequency range and a high signal-to-noise ratio [48, 68, 69], and has recently been transferred to terahertz communications [70].

To exploit the tremendous application potential of optoelectronic T-wave processing, monolithic co-integration of photonic devices and T-wave transmitters and receivers is of vital importance. From the technology side, however, optoelectronic T-wave transmitters and receivers are still rather complex, relying on high-speed photodiodes [21, 71, 72] or photoconductors [48, 73, 74] that require dedicated III–V semiconductor substrates (obtained, for

example, through low-temperature growth of InGaAs/InAlAs multilayer structures [48]) and which are not amenable to large-scale photonic integration. This not only hampers the co-integration of T-wave transmitter and receiver circuitry on a common chip, but also hinders the exploitation of highly developed photonic integration platforms to build advanced optoelectronic T-wave systems that combine photonic signal processing with optoelectronic frequency conversion on a common chip.

In this article, we demonstrate an approach that allows the integration of T-wave transmitters and receivers directly on the silicon photonic platform, thereby exploiting the outstanding technical maturity, scalability and the comprehensive device portfolio [75, 76] of this material system. The approach exploits internal photoemission at the metal–semiconductor interfaces of plasmonic structures [4], which can be directly integrated into widely used silicon-on-insulator (SOI) waveguides. Our experiments show that these plasmonic internal-photoemission detectors (PIPEDs) are not only suited for photomixing at the T-wave transmitter, but also lend themselves to highly sensitive optoelectronic reception. In a proof-of-concept experiment, we monolithically co-integrated a PIPED transmitter and a PIPED receiver on a common silicon photonic chip and used them to measure the complex transfer function of an integrated T-wave transmission line. In this context, we also developed and experimentally verified a mathematical model of optoelectronic T-wave conversion that allows us to quantitatively describe T-wave generation and detection over a wide range of frequencies.

6.2   Silicon-Plasmonic T-wave Systems

The vision of an integrated silicon-plasmonic T-wave system is illustrated in Figure 20 using a wireless high-speed transceiver as an exemplary application case. The system combines a T-wave transmitter, a T-wave receiver and a variety of other silicon photonic devices [75, 77] such as phase shifters [78-80] or high-performance modulators [81-84] on a common substrate. C.w. lasers are coupled to the chip using photonic wire bonds [85], and electrical circuits such as field-programmable gate arrays (FPGA), digital-to-analog and analog-to-digital converters (DAC and ADC) are used to drive the modulators and to further process the received signals. T-wave generation is accomplished by
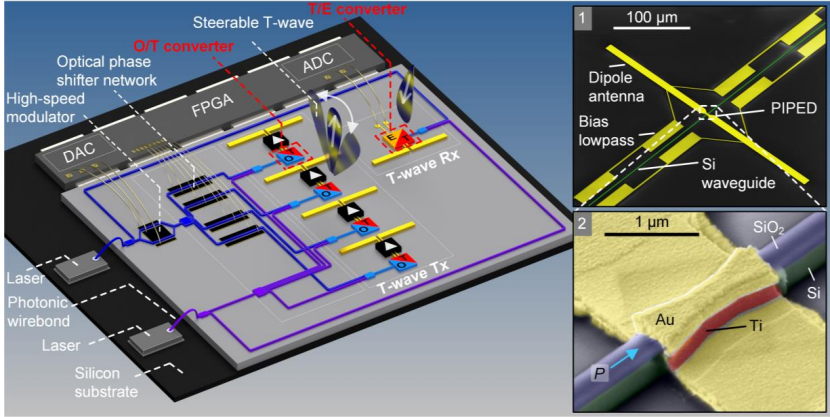
Figure 20. Vision of an integrated silicon-plasmonic T-wave wireless transceiver that exploits optoelectronic signal processing both at the transmitter and at the receiver. The system combines optical-to-T-wave (O/T) and T-wave-to-electronic (T/E) converters with advanced silicon photonic devices such as phase shifters or high-performance modulators. Continuous-wave lasers are coupled to the chip using photonic wire bonds [85], and electrical circuits such as field-programmable gate arrays (FPGAs), DACs and ADCs are used to drive the modulators and to process the received signals. The transmitter comprises a T-wave antenna array, fed by an array of O/T converters that are driven by a series of optical signals. The phases of the optical and hence the T-wave signals can be precisely defined by an electrically driven optical phase shifter network, thereby enabling broadband beam steering. Optionally, additional T-wave amplifiers could boost the power of the generated signals. Coherent detection of the T-wave signal relies on a pair of optical carriers, the power beat of which serves as a LO for T/E conversion. **Inset 1**: false-color scanning electron microscopy (SEM) image of an O/T or T/E converter. The devices rely on PIPEDs [4] coupled to dipole antennas. The PIPEDs are fed through silicon photonic waveguides and biased via low-pass structures that are directly connected to the arms of the antennas. **Inset 2**: detailed view of a fabricated PIPED. The device consists of a narrow silicon nanowire waveguide that is combined with overlays of Au and Ti to form an ultrasmall plasmonic structure with two metal-semiconductor interfaces. Optical power P is fed to the PIPED by a silicon photonic waveguide.

photomixing the modulated optical signals with an optical c.w. tone in a high-speed antenna-coupled PIPED acting as an optical-to-T-wave (O/T) converter [4]. Large-scale monolithic integration of advanced silicon photonic devices with O/T converters opens up rich opportunities for advanced T-wave signal processing. Note that, for integrated T-wave systems, the generation and detection of c.w. signals offers various advantages over pulsed operation (see Methods in [6] for details). In the example, the transmitter comprises a T-wave

antenna array, fed by an array of O/T converters that are driven by a series of optical signals. The phases of the optical signals and hence those of the T-waves can be precisely defined by an electrically driven optical phase shifter network [78, 79], thereby enabling broadband beam steering and shaping. Optionally, integrated T-wave amplifiers can be used to boost the T-wave signals [86]. At the receiver, optoelectronic down-conversion (T/E-conversion) is used for coherent detection of the T-wave signal, using the power beat of two optical waves as a LO. For O/T and T/E conversion, the concept relies on PIPEDs that are coupled to dipole antennas (Inset 1 of Figure 20).

The PIPEDs are fed through silicon photonic waveguides and electrically biased by dedicated low-pass structures that are directly connected to the arms of the dipole antennas. A more detailed view of a PIPED is shown in Inset 2 of Figure 20. The device consists of a narrow silicon nanowire waveguide that is combined with overlays of gold (Au) and titanium (Ti) to form an ultra-small plasmonic structure with two metal–semiconductor interfaces. Note that the PIPED concept does not rely on the use of gold as a plasmonic material – this was chosen only for ease of fabrication in the current experiment. When combined with large-scale silicon photonic circuits, gold-free designs may be used that allow processing in a state-of-the-art CMOS line [87].

6.3   PIPED for Optoelectronic T-wave Processing

The PIPED concept is illustrated and explained in Figure 21. Figure 21 (a) presents a schematic cross-section of the device. The Si nanowire waveguide core is contacted by an Au layer on the left and by a Ti layer on the right (for details of the fabrication see ref. [4]). To drive the device, light at infrared telecommunication wavelengths ($\lambda \approx 1.5\,\mu m$) is coupled to the Si waveguide core, leading to excitation of surface plasmon polaritons (SPPs) both at the Au-Si and Si–Ti interfaces. The associated energy levels are sketched in Figure 21 (b) for a forward bias voltage $U > 0$, which is counted positive from the Au to Ti electrode. Free-carrier absorption generates hot electrons in the titanium with carrier energies above the Fermi level $W_{F,Ti}$. An equivalent effect occurs for holes at the Au–Si interface (the relative magnitude of the two contributions is currently under investigation). The hot electrons and holes have
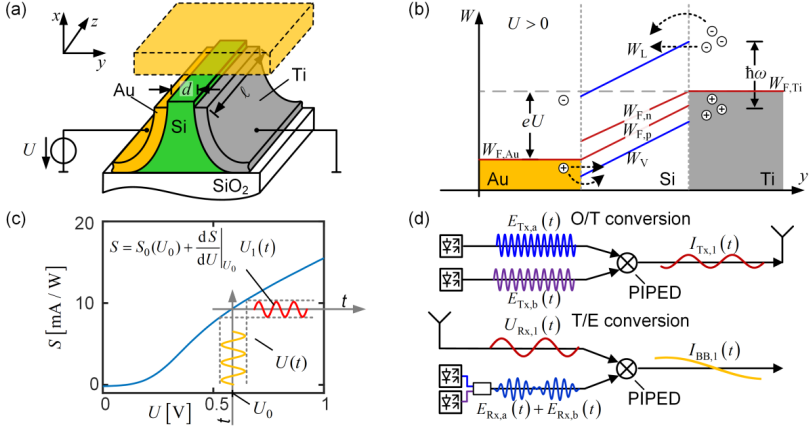
Figure 21. Operating principle of PIPED-based T-wave transmitters and coherent receivers. **(a)** Schematic of the PIPED, consisting of a silicon core with Au and Ti sidewalls. SPPs propagating along the z direction are mostly absorbed in the Ti due to the large imaginary part of its complex electric permittivity ε_r. A voltage U is applied between the Au and Ti electrodes. **(b)** Band diagram of the Au–Si–Ti junction. W_C and W_V denote the edges of the conduction and valence bands in the silicon core, and the Fermi levels in Au and Ti are denoted $W_{F,Au}$ and $W_{F,Ti}$, respectively. An applied voltage $U > 0$ and the injection of carriers due to absorption of light lead to an interband non-equilibrium, which is described by the separation of the quasi-Fermi levels of conduction-band electrons and valence-band holes, denoted $W_{F,n}$ and $W_{F,p}$, respectively. Photons absorbed in the Ti layer excite hot electrons, which can overcome or tunnel through the Schottky potential barrier, leading to a photocurrent $I = S \times P$. An equivalent effect occurs for holes at the Au–Si interface. The relative magnitude of the two contributions is still under investigation. The carrier emission probability into the Si waveguide core and therefore the sensitivity $S(U)$ can be increased ($U > 0$) or decreased ($U < 0$) by varying bias voltage U. **(c)** Measured sensitivity $S(U)$ of the PIPED in dependence on applied voltage U. In the vicinity of a bias voltage, the sensitivity $S(U)$ can be linearized. **(d)** In essence, the PIPED can be used as a mixer that multiplies two signals to generate a waveform at the difference frequency. Top: when used for photomixing (O/T conversion) at the T-wave transmitter (Tx), the PIPED acts as a power detector, and the output photocurrent corresponds to the product of two time-dependent optical signals $E_{Tx,a}(t)$ and $E_{Tx,b}(t)$. Bottom: when used for optoelectronic down-conversion (T/E conversion) at the T-wave receiver (Rx), the PIPED combines two functionalities, namely the generation of a terahertz LO from two optical carriers $E_{Tx,a}(t)$ and $E_{Tx,b}(t)$ and the down-conversion of the received T-wave to the baseband. To this end, the PIPED is fed by a superposition of two unmodulated optical carriers, oscillating at frequencies $f_{Rx,a}$ and $f_{Rx,b}$, while a time-dependent voltage $U_{Rx,1}(t)$ modulates the device sensitivity. The PIPED photocurrent is then given by the product of the time-variant sensitivity with the time-variant optical power $P_{Rx}(t)$.

an increased probability to cross the $d = 100$ nm wide Si barrier, leading to a photocurrent I from the Au to the Ti side. The photocurrent I depends linearly on optical power P with a sensitivity (or responsivity) $S = I / P$. Note that, in contrast to conventional photodiodes, the measured sensitivity $S(U)$ of the PIPED depends strongly on the applied voltage U, as shown in Figure 21 (c). This voltage dependence is a key aspect for efficient optoelectronic T/E conversion as it allows mixing of an oscillating optical input power with a time-dependent voltage applied to the PIPED contacts. For a forward bias $U > 0$, the band edges tilt inside the Si core (Figure 21 (b)), which reduces the effective width of the potential barrier such that the sensitivity S increases with U. For a reverse bias $U < 0$, the carrier emission probability is small, and so the photocurrent remains small. The strong absorption of the SPP allows junction lengths ℓ of less than 1 μm and device capacitances smaller than $C = 1$ fF (ref. [4]). With a load resistance of $R = 50\,\Omega$, this would lead to an RC limiting frequency of 3 THz.

In the current device designs, this limitation is not relevant because the speed is limited by the carrier transit times, for which we estimate 1 ps for electrons and 1.5 ps for holes. This estimate is based on a voltage drop of $U = 0.5$ V within the 100-nm-wide barrier, leading to drift velocities of $10^7\,\mathrm{cms}^{-1}$ for electrons and $6.5 \cdot 10^5\,\mathrm{cms}^{-1}$ for holes [88], close to the respective saturation velocities. This limits the bandwidth to ~0.44 THz assuming dominating electron transport and to ~0.29 THz in case hole transport dominates [89]. The fast device response makes the PIPED an excellent candidate for T-wave generation and reception at frequencies of 1 THz and above. It is worth noting that the sensitivity of the PIPED can be increased beyond the values depicted in Figure 21 (c) by applying larger bias voltages. We have previously demonstrated sensitivities of up to 0.12 A W^{-1} (ref. [4]) - the highest value so far demonstrated for photodetectors based on internal photoemission. This sensitivity is still below the 0.2–0.3 A W^{-1} typically achieved for the uni-travelling-carrier photodiodes (UTC-PD) designed to operate at frequencies around 0.3 THz (refs. [21-23]), and we expect that further improvements are possible by optimizing the materials and geometries of the PIPED. Note that the sensitivity of the photodetector is not of utmost importance for the generation and detection of T-wave signals, because a smaller sensitivity can be compensated by launching a higher optical power.

For O/T conversion, the transmitter (Tx) essentially acts as a mixer multiplying two time-dependent optical signals $E_{\text{Tx,a}}(t)$ and $E_{\text{Tx,b}}(t)$ to produce a photocurrent $I_{\text{Tx}}(t)$ that corresponds to the difference-frequency waveform, (Figure 21 (d), top). In the following, we only give a short mathematical description of photomixing and optoelectronic down-conversion in the PIPED. A rigorous analysis is provided in Supplementary Section 1 in [6]. We assume that the optical signal $E_{\text{Tx,a}}(t)$ oscillates at angular frequency $\omega_{\text{Tx,a}}$ and carries an amplitude modulation $\hat{E}_{\text{Tx,a}}(t)$ and/or a phase modulation $\varphi_{\text{Tx,a}}(t)$, whereas the optical signal $E_{\text{Tx,b}}(t)$ is simply a c.w. carrier with constant amplitude $\hat{E}_{\text{Tx,b}}$, frequency $\omega_{\text{Tx,b}}$ and phase $\varphi_{\text{Tx,b}}$:

$$
\begin{aligned}
E_{\text{Tx,a}}(t) &= \hat{E}_{\text{Tx,a}}(t)\cos\left(\omega_{\text{Tx,a}}t + \varphi_{\text{Tx,a}}(t)\right), \\
E_{\text{Tx,b}}(t) &= \hat{E}_{\text{Tx,b}}\cos\left(\omega_{\text{Tx,b}}t + \varphi_{\text{Tx,b}}\right).
\end{aligned}
\tag{41}
$$

The optical power $P_{\text{Tx}}(t)$ then oscillates at the difference frequency $\omega_{\text{Tx,THz}} = |\omega_{\text{Tx,a}} - \omega_{\text{Tx,b}}|$,

$$
P_{\text{Tx,l}}(t) = \hat{P}_{\text{Tx,l}}(t)\cos\left(\omega_{\text{Tx,THz}}t + \varphi_{\text{Tx,THz}}(t)\right),
\tag{42}
$$

where the amplitude $\hat{P}_{\text{Tx,l}}(t)$ and phase $\varphi_{\text{Tx,THz}}(t)$ of the oscillation are directly linked to the normalized amplitude and to the phase of the optical wave:

$$
\hat{P}_{\text{Tx,l}}(t) = \hat{E}_{\text{Tx,a}}(t)\hat{E}_{\text{Tx,b}}, \; \varphi_{\text{Tx,THz}}(t) = \varphi_{\text{Tx,a}}(t) - \varphi_{\text{Tx,b}}.
\tag{43}
$$

When detected by the PIPED (sensitivity S_{Tx}), this leads to an oscillating component in the photocurrent $I_{\text{Tx}}(t)$, featuring the same frequency and the same phase as the optical power oscillation:

$$
I_{\text{Tx,l}}(t) = \hat{I}_{\text{Tx,l}}(t)\cos\left(\omega_{\text{Tx,THz}}t + \varphi_{\text{Tx,THz}}(t)\right),
\tag{44}
$$

where

$$
\hat{I}_{\text{Tx,l}}(t) = S_{\text{Tx}}\hat{E}_{\text{Tx,a}}(t)\hat{E}_{\text{Tx,b}}
\tag{45}
$$

Hence, any modulation of the amplitude $\hat{E}_{\text{Tx,a}}(t)$ or phase $\varphi_{\text{Tx,a}}(t)$ of the optical signal translates directly into an amplitude and phase modulation of the T-wave. The amplitude of the emitted terahertz field is linearly connected to the

sensitivity S_{Tx} and can be increased by applying a larger forward bias to the PIPED. Exploiting this concept, broadband high-quality terahertz signals can be generated by using widely available optical communication equipment. The terahertz field is radiated by an antenna or coupled to a transmission line.

Similarly, PIPEDs can be used for T/E conversion in the T-wave receiver (Rx) (Figure 21 (d), bottom). In this case, the device combines two functionalities, namely the generation of a terahertz LO from two optical carriers and the down-conversion of the received T-wave to the baseband. To this end, the PIPED is fed by a superposition of two unmodulated optical tones, oscillating at frequencies $\omega_{Rx,a}$ and $\omega_{Rx,b}$:

$$
\begin{aligned}
E_{Rx,a}(t) &= \hat{E}_{Rx,a} \cos\left(\omega_{Rx,a} t + \varphi_{Rx,a}\right), \\
E_{Rx,b}(t) &= \hat{E}_{Rx,b} \cos\left(\omega_{Rx,b} t + \varphi_{Rx,b}\right).
\end{aligned}
\tag{46}
$$

This leads to an oscillating power at frequency $\omega_{Rx,THz} = \left|\omega_{Rx,a} - \omega_{Rx,b}\right|$ with phase $\varphi_{Rx,THz}$:

$$
P_{Rx,l}(t) = \hat{P}_{Rx,l} \cos\left(\omega_{Rx,THz} t + \varphi_{Rx,THz}\right),
\tag{47}
$$

where

$$
\hat{P}_{Rx,l} = \hat{E}_{Rx,a} \hat{E}_{Rx,b}, \quad \varphi_{Rx,THz} = \varphi_{Rx,a} - \varphi_{Rx,b}.
\tag{48}
$$

At the same time, the PIPED is biased with a DC voltage $U_{Rx,0}$, which is superimposed by the time-variant terahertz signal $U_{Rx,l}(t)$ generated by the terahertz antenna. The overall time-dependent voltage applied to the PIPED hence reads

$$
U_{Rx}(t) = U_{Rx,0} + U_{Rx,l}(t),
\tag{49}
$$

where

$$
U_{Rx,l}(t) = \hat{U}_{Rx,l}(t) \cos\left(\omega_{Tx,THz} t + \varphi_{Tx,THz}(t) - \varphi_{TxRx}\right).
\tag{50}
$$

In this relation, the phase at the receiver depends on the phase delay φ_{TxRx} that the T-wave experiences when propagating from the Tx to the Rx. Due to the voltage-dependent PIPED sensitivity, the time-varying voltage $U_{Rx,l}(t)$ leads to

a temporal variation of the sensitivity $S_{\mathrm{Rx}}\big(U_{\mathrm{Rx,1}}(t)\big)$, oscillating at frequency $\omega_{\mathrm{Tx,THz}}$ of the incident T-wave. The PIPED photocurrent is given by the product of the time-variant sensitivity with the time-variant optical power $P_{\mathrm{Rx,1}}(t)$. For the case of homodyne detection, $\omega_{\mathrm{Rx,THz}} = \omega_{\mathrm{Tx,THz}} = \omega_{\mathrm{THz}}$, the baseband current at the output of the Rx PIPED is given by

$$
\begin{aligned}
I_{\mathrm{BB}}(t) &= I_{\mathrm{BB,0}} + I_{\mathrm{BB,1}}(t) \\
&= I_{\mathrm{BB,0}} + \hat{I}_{\mathrm{BB,1}}(t)\cos\big(\varphi_{\mathrm{BB}}(t)\big),
\end{aligned}
\tag{51}
$$

where the amplitude $\hat{I}_{\mathrm{BB,1}}(t)$ of the time-variant part of the baseband photocurrent and the associated time-variant phase $\varphi_{\mathrm{BB}}(t)$ are connected to the amplitude and phase of the time-variant terahertz signal $U_{\mathrm{Rx,1}}(t)$:

$$
\hat{I}_{\mathrm{BB,1}}(t) = \tfrac{1}{2}\left.\frac{\mathrm{d}\,S_{\mathrm{Rx}}}{\mathrm{d}\,U_{\mathrm{Rx}}}\right|_{U_{\mathrm{Rx,0}}} \hat{P}_{\mathrm{Rx,1}}\hat{U}_{\mathrm{Rx,1}}(t) \;,
$$

$$
\varphi_{\mathrm{BB}}(t) = \varphi_{\mathrm{Tx,THz}}(t) - \varphi_{\mathrm{Rx,THz}} + \varphi_{\mathrm{TxRx}}\,.
\tag{52}
$$

For sensitive detection, the slope $\mathrm{d}\,S_{\mathrm{Rx}}/\mathrm{d}\,U_{\mathrm{Rx}}\big|_{U_{\mathrm{Rx,0}}}$ of the sensitivity as a function of voltage has to be maximized such that small variations of the terahertz voltage translate into large variations of the baseband photocurrent amplitude. The phase $\varphi_{\mathrm{BB}}(t)$ may be properly adjusted with the variable time delay τ_{opt} (Figure 22 (a)). A more convenient alternative measurement technique is explained in the Methods of [6].

6.4   Demonstration of T-wave Generation and Detection

For an experimental demonstration of the PIPED performance in O/T and T/E conversion, we first characterized the Tx and Rx separately. To this end, we fabricated a PIPED that was connected to an on-chip dipole antenna as shown in Inset 1 of Figure 20. To supply a bias voltage $U_{\mathrm{Tx,0}}$ to the PIPED, we used bias lines equipped with terahertz chokes that prevent leakage of terahertz signals from the antenna; see Supplementary Section 3 in [6] for details of the antenna design. Note that the dipole antenna is not optimum for transmitting

and receiving terahertz power over a broad frequency range. We still decided to use this antenna concept due to its small footprint, which allowed us to densely integrate a large number of PIPEDs on a test chip. The dipole antennas may be replaced by bigger spiral or bow-tie structures in future devices. To measure the device performance, we used the set-up depicted in Figure 22 (a), where the Tx and Rx are driven by the same lasers for homodyne detection. To increase the sensitivity of T-wave detection, we used a modulated bias voltage $U_{Tx,0}$ that leads to a modulated terahertz power and helps in detecting the received T-wave with a lock-in amplifier. The T-wave is transmitted via a silicon lens and redirected to the lensed Rx antenna by an off-axis parabolic mirror.
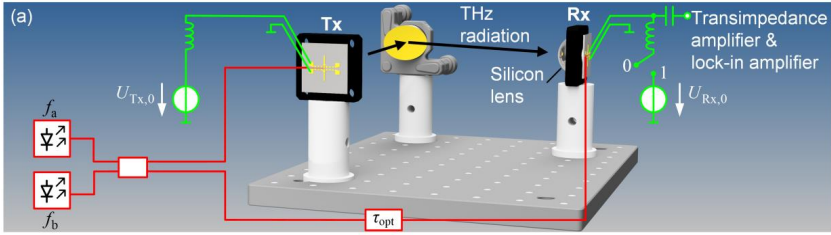
Figure 22. Demonstration of PIPED performance for O/T and T/E conversion. **(a)** Experimental set-up. To measure the PIPED Tx and Rx separately, the respective other component is replaced by a commercially available device in a free-space set-up. Light from two lasers with frequencies f_a and f_b is superimposed to obtain an optical power beat at $f_{THz} = |f_a - f_b|$, which generates a T-wave in the PIPED Tx or acts as a LO for homodyne detection in the PIPED Rx. When using the PIPED at the Tx, the device is biased by a DC voltage $U_{Tx,0}$, which is modulated by an AC signal for lock-in detection. The bias is applied to the PIPED Tx via dedicated on-chip low-pass structures (terahertz chokes), see Inset 1 of Figure 20. The same structures are used at the PIPED Rx, along with an additional bias-T in the feed circuit that allows the DC bias to be separated from the AC lock-in signal. The AC signal is fed to a transimpedance amplifier and detected by a lock-in amplifier. (Figure continues on the next page.)
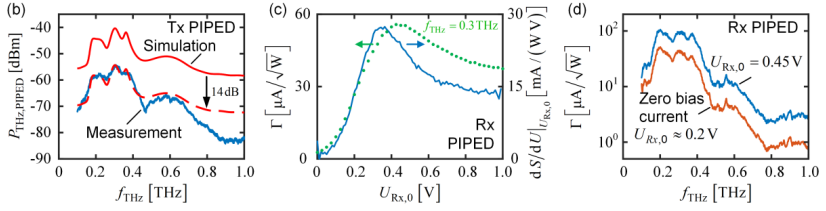
When operating the PIPED as a T-wave Rx, we either apply a defined bias voltage $U_{Rx,0}$ by connecting an external voltage source (switch position 1), or we leave the bias contacts open (switch position 0, 'zero bias current'), which leads to a build-up of an internal forward bias of $U_{Tx,0}$ upon illumination of the PIPED. **(b)** PIPED as a T-wave Tx. The radiated power (blue) was determined by a commercial reference Rx (Toptica, EK – 000725). The Tx antenna and terahertz chokes were designed for a resonance frequency of 0.385 THz (see Supplementary Section 3 in [6]). The peaks at lower frequencies are caused by unwanted resonances on the choke structures. For comparison, we also plot the simulated terahertz power (solid red) emitted by the PIPED via the silicon substrate and an adjacent silicon lens. The shape of the measured frequency characteristic coincides reasonably well with its simulated counterpart, while the absolute power levels obtained from the measurement are ~14 dB below the simulated values. We attribute this to losses in the set-up and to imperfect coupling of the PIPED terahertz beam to the reference Rx. For better comparison of the shapes, we shifted the simulated characteristic by 14 dB (dashed red) to coincide with the measurement. Details of the simulation are provided in the Methods and Supplementary Section 9 in [6]. **(c)** PIPED as T-wave Rx. Measured conversion factor $\Gamma = \hat{I}_{BB,1}/\sqrt{P_{THz}}$ (green) as a function of Rx bias voltage $U_{Rx,0}$. The bias-dependent sensitivity slope $dS/dU|_{U_{Rx,0}}$ (blue) features a maximum that deviates slightly from the maximum of the conversion factor Γ at a carrier frequency of 0.3 THz. We attribute this effect to the voltage dependence of the transit time. **(d)** Conversion factor Γ of the Rx as a function of the terahertz frequency both for zero-bias-current operation ($U_{Rx,0} \approx 0.2\,\mathrm{V}$) and for an externally applied DC bias of $U_{Rx,0} = 0.45\,\mathrm{V}$, which leads to the maximum conversion factor in (c).

To characterize the PIPED Tx performance, we determined the emitted power as a function of frequency by comparison to a commercially available terahertz reference Tx (Toptica, EK – 000724) (see Methods in [6]). To ensure traceability, the emitted power $P_{THz,\,ref}$ of the reference Tx was measured using a calibrated pyroelectric thin-film sensor. We then detected the emission of the PIPED and the reference Tx with a commercially available photoconductive terahertz receiver (Toptica, EK – 000725). The terahertz power $P_{THz,PIPED}$ generated by the PIPED Tx leads to a current $I_{Rx,PIPED}$ in the reference receiver,

while the emission of the reference Tx generates a receiver current $I_{\text{Rx,ref}}$. The PIPED terahertz output power can then be estimated by

$$P_{\text{THz,PIPED}} = P_{\text{THz, ref}} \left(I_{\text{Rx,PIPED}} / I_{\text{Rx,ref}} \right)^2. \tag{53}$$

The frequency dependence of the terahertz power emitted by the PIPED is shown in Figure 22 (b). The PIPED is capable of generating radiation at frequencies of up to 1 THz. The measured transmitted power $P_{\text{THz,PIPED}}$ obtained according to Equation (53) is depicted as a blue solid line.

For comparison, we perform a simulation of the terahertz power that the PIPED emits into free space via the silicon substrate and an adjacent silicon lens; the results are plotted as a solid red line. Details of the simulation are provided in the Methods and Supplementary Section 9 in [6]. The shapes of the measured and simulated frequency characteristics coincide reasonably well, but the absolute power levels obtained from the measurement are ~14 dB below the simulated values. We attribute this effect to losses in the set-up and to imperfect coupling of the PIPED terahertz beam to the reference Rx. The peaks in the simulated and measured frequency characteristic are caused by antenna and bias line resonances. The roll-off at larger frequencies is predominantly caused by the carrier transit time in the 100 nm-wide silicon core of the PIPED (Supplementary Section 9 in [6]). This width can be further reduced [4], for example to 75 nm, which would decrease the transit time accordingly. To avoid irreversible damage to the device in our Tx experiments, we chose to limit the optical power such that the mean photocurrent did not exceed 50 µA. The measured radiated output power exhibits a maximum of approximately –55 dBm at 0.3 THz. Note that this value, as well as its simulated counterpart of approximately –40 dBm, are still well below the power levels that can be achieved by state-of-the-art discrete terahertz sources. As an example, commercially available electronic multipliers (www.vadiodes.com) permit output powers of tens of milliwatts at 0.3 THz when operated with input powers of hundreds of milliwatts at 150 GHz, and similar levels can be achieved with IMPATT diodes (www.terasense.com). However, these devices are limited to certain frequency bands or even to single-frequency operation for the case of IMPATT diodes. Photomixing in high-speed UTC-PDs on III–V substrates allows for broadband operation, with power levels of up to 1.2 mW at 0.3 THz

achieved for a pair of devices operated in parallel [21]. The power level of our commercial terahertz reference Tx is −19 dBm at 0.3 THz, measured by a calibrated pyroelectric detector (Supplementary Section 2 in [6]). In comparison to these established terahertz generators, our current PIPED devices are still inferior in terms of maximum output power. Note, however, that the PIPED approach exploits the intrinsic scalability advantages of highly mature silicon photonic integration and is hence perfectly suited for integrated T-wave signal processing systems that combine the T-wave Tx and Rx on a common chip. These systems can be operated at very low power levels as shown in the next section. In addition, the output power of PIPED systems can be further increased. For instance, the maximum applied current of 50 µA is still well below the actual damage threshold of the PIPED, and we expect that much higher operating currents can be safely applied provided that the devices are thermally connected to a proper heat sink (Supplementary Section 5 in [6]). Moreover, the output power can be increased by connecting an array of devices to a single antenna or transmission line (Supplementary Section 12 in [6]). Exploiting the coherent superposition of the individual terahertz currents, an array of n synchronously pumped PIPEDs connected in parallel would increase the terahertz output power of the Tx by a factor of n^2. The small PIPED length of roughly 1 µm, the small capacitance below 1 fF, and the large resistance in excess of 10 kΩ would easily allow the connection of $n = 10$ PIPEDs to the antenna feed point, leading to a power gain of 20 dB. Moreover, the intrinsic scalability of silicon photonics allows large-scale antenna arrays to be built on a single chip, which could further increase the output power as well as the directivity of the emitted terahertz power.

To evaluate the PIPED Rx performance, we used a commercially available Tx (Toptica, EK − 000724). According to Equation (52), the baseband photocurrent amplitude $\hat{I}_{BB,1}$ depends linearly on the amplitude $\hat{U}_{Rx,1}$ of the terahertz voltage, which is proportional to the square root of terahertz power P_{THz}. As a metric for Rx sensitivity, we can hence define the ratio

$$\Gamma = \frac{\hat{I}_{BB,1}}{\sqrt{P_{THz}}} \propto \left. \frac{dS_{Rx}}{dU_{Rx}} \right|_{U_{Rx,0}} \hat{P}_{Rx,1}, \tag{54}$$

which describes the conversion factor from the terahertz signal to the baseband photocurrent. The linear relationship between $\hat{I}_{\mathrm{BB,l}}$ and $\sqrt{P_{\mathrm{THz}}}$ was confirmed experimentally (Supplementary Section 4 in [6]). The conversion factor depends on the sensitivity slope $\mathrm{d}\,S_{\mathrm{Rx}}/\mathrm{d}\,U_{\mathrm{Rx}}|_{U_{\mathrm{Rx,0}}}$, see Equation (52). This was experimentally confirmed by measuring the conversion factor Γ in dependence on the bias voltage $U_{\mathrm{Rx,0}}$ at a frequency of 0.3 THz. The result is depicted in Figure 22 (c) (green, left axis) along with the slope of the sensitivity (blue, right axis) derived from the static $S(U)$ characteristic in Figure 21 (c). The two curves are in fair agreement. We also verified that the conversion factor depends linearly on the incident optical power $\hat{P}_{\mathrm{Rx,l}}$ (Supplementary Section 4 in [6]). These findings confirm the validity of our PIPED model used to describe the Rx.

Similarly, we demonstrate the ability of a PIPED to perform broadband T/E conversion at the Rx. To this end, we measured the conversion factor Γ as a function of terahertz frequency for two cases. In a first measurement, we left the bias contacts open, which corresponds to switch position 0 in Figure 22 (a) (zero-bias-current operation). This leads to the build-up of an internal forward bias of $U_{\mathrm{Rx,0}} \approx 0.2\,\mathrm{V}$ when illuminating the PIPED (Supplementary Section 6 in [6]). For the second measurement, we turned the switch in Figure 22 (a) to position 1 and connected an external DC voltage source $U_{\mathrm{Rx,0}} = 0.45\,\mathrm{V}$. The results of the measured conversion efficiencies Γ are shown Figure 22 (d). The PIPED is able to receive radiation at frequencies up to 1 THz and beyond. At a frequency of 0.3 THz, the devices exhibit a conversion factor of $95\,\mu\mathrm{A}/\sqrt{\mathrm{W}}$ for bias voltages of $U_{\mathrm{Rx,0}} = 0.45\,\mathrm{V}$ and $44\,\mu\mathrm{A}/\sqrt{\mathrm{W}}$ for zero bias current, that is, internal bias voltages of $U_{\mathrm{Rx,0}} \approx 0.2\,\mathrm{V}$. As expected from Figure 22 (c), the bias voltage of 0.45 V leads to a larger slope $\mathrm{d}\,S_{\mathrm{Rx}}/\mathrm{d}\,U_{\mathrm{Rx}}|_{U_{\mathrm{Rx,0}}}$ of the sensitivity and hence to a higher conversion factor as compared to the 0.2 V bias. The resonances in Figure 22 (d) and the drop of the conversion factor for larger frequencies are caused by the frequency response of the antenna, the PIPED and the bias lines.

If we assume a $50\,\Omega$ load, the conversion factor Γ can be translated to a conversion gain of $\frac{1}{2}\Gamma^2 \times 50\,\Omega$, which corresponds to -73 dB for zero-bias-current operation and -66 dB for a bias voltage of 0.45 V. This is clearly below

the conversion gain of -8.5 dB that can be achieved by state-of-the-art electronic subharmonic mixers (www.vadiodes.com) operating at 300 GHz with a 150 GHz electronic local oscillator. Note, however, that the PIPED T-wave receivers can be tuned across a large range of operation frequencies that exceeds that of discrete mixers, which are usually limited to the specific transmission band of the waveguides needed to package the discrete devices and to block unwanted frequency components. Regarding optoelectronic down-conversion of T-wave signals, InP-based UTC-PDs have been exploited [22, 90, 91], leading to a conversion gain of –30 dB at a frequency of 0.3 THz (ref. [22]). However, down-conversion in UTC-PDs relies on the dynamic capacitance associated with accumulation of carriers in the absorption region and is hence only efficient for very high optical powers (for example, 80 mW) [22]. This leads to large DC currents and hence a strong noise background in the baseband current (for example, $28 \mathrm{nA}/\sqrt{\mathrm{Hz}}$,); see Methods in [6] for details. Hence, despite the large conversion gain, the baseband signal-to-noise power ratio (SNR) of UTC-PD T-wave receivers is limited, and levels of 57 dB Hz^{-1} have been demonstrated at a frequency of 0.3 THz for the devices of ref. [91]. In contrast to that, PIPEDs exploit an intrinsically voltage-dependent sensitivity (Figure 22 (c)) and thus require only relatively small optical power levels. For a PIPED operated with zero bias current, we measured a noise current of $9 \mathrm{pA}/\sqrt{\mathrm{Hz}}$ and a SNR of $85 \mathrm{dB}\,\mathrm{Hz}^{-1}$ at 0.3 THz. These figures compare very well to the noise current of $15 \mathrm{pA}/\sqrt{\mathrm{Hz}}$ and SNR of $91 \mathrm{dB}\,\mathrm{Hz}^{-1}$ reported for state-of-the-art photoconductors [48] as well as to the $18 \mathrm{pA}/\sqrt{\mathrm{Hz}}$ and $84 \mathrm{pA}/\sqrt{\mathrm{Hz}}$ that we measured when replacing the PIPED Rx with photoconductive reference Rx. The conversion gain of the reference Rx is – 67 dB at 0.3 THz (Supplementary Figure 1 (b) in [6]) and hence is comparable to that of the PIPED Rx.

Based on these findings, we conclude that a PIPED T-wave Rx can already provide conversion factors and SNR levels that are comparable to those of conventional III–V devices while offering an ultra-compact footprint. It is also worth noting that the PIPEDs reported here are first-generation devices that still feature large potential for further performance improvements. As an example, using parallel arrays of n PIPEDs connected to a single Rx antenna would allow boosting the power conversion factor in proportion to n^2, hence allowing improvements of the conversion gain by one to two orders of magnitude

(Supplementary Section 12 in [6]). Note also that the load resistance of 50 Ω assumed for translating the conversion factor Γ into a conversion gain is not the optimum choice: the PIPED itself features a large output impedance well above 10 kΩ, and the conversion gain would hence increase if load resistances larger than 50 Ω could be used. Given the ultrasmall capacitance of the PIPED (Supplementary Section 9 in [6]), this would be possible without imposing significant RC limitations to the bandwidth of the device. These aspects indicate that the PIPED-based T-wave Rx has the clear potential to unlock a wide variety of applications that rely on optoelectronic T-wave signal processing in compact integrated systems.

The SNR achieved with our PIPED Rx is sufficient for many spectroscopic or diagnostic applications, where lock-in detection and large integration times lead to a small detection bandwidth. Application of Rx PIPEDs in high-speed terahertz communication would still require higher SNR, dictated by the large bandwidth of the data signal. Assuming that a viable transmission link requires a signal power that is at least 10 dB larger than the noise power in a 10 GHz signal bandwidth, the SNR needs to be larger than 110 dB Hz^{-1}. By using a high-power transmitter, for example, an UTC-PD [21], or by using a T-wave amplifier [86], T-wave powers of more than 0 dBm can be achieved, which would increase the SNR to more than 100 dB Hz^{-1}. Combining these transmitters with parallel arrays of PIPEDs connected to the same Rx antenna would lead to SNR values suitable for wireless communication links.

6.5   Monolithically Integrated T-wave System

To demonstrate the technological advantages of the PIPED concept, we monolithically integrate arrays of PIPED Tx and PIPED Rx on a common silicon chip. Tx and Rx are coupled by short T-wave transmission lines having various lengths L between 10 μm and ~1 mm. An SEM image of such a Tx–Rx pair is displayed in Figure 23 (a). A 1-μm-wide gap in the middle of the T-wave transmission line acts as a DC block to decouple the bias voltages of the Tx and Rx. The insets in Figure 23 (a) show magnified pictures of the Tx, the Rx, and the gap. In the experiment, we used the PIPED Tx and Rx to measure the amplitude and phase transfer characteristics of the transmission line at terahertz frequencies; the associated set-up is shown in Figure 23 (b). Optical signals are

fed to the Tx and the Rx PIPED by separate silicon photonic waveguides equipped with grating couplers. As in the previous experiments, the Tx and Rx are fed with two optical spectral lines f_a and f_b at varying frequency separations. The Tx is biased in the forward direction $U_{Tx,0} > 0$ and feeds the T-wave transmission line. At the end of the transmission line, the Rx PIPED acts as a homodyne receiver to coherently detect the T-wave signal and to down-convert it to a baseband current $I_{BB,1}$, (51) and (52). The Rx PIPED is operated under zero-bias-current conditions to minimize noise. Because the PIPED features a large impedance at terahertz frequencies, all transmission lines are essentially terminated by open circuits at both ends and thus act as T-wave resonators, for which the resonance frequencies are dictated by the respective geometric length L. At the Rx, we measure the baseband photocurrent $I_{BB,1}$. By sweeping the Tx frequency, we can separate the amplitude $I_{BB,1}$ from the phase φ_{BB} (see Methods in [6] for details). To quantify the transmission line transfer characteristics, we used the complex transfer impedance $\underline{Z}_{21} = |\underline{Z}_{21}| e^{j\varphi_{21}}$, which is defined as the ratio of the complex T-wave voltage amplitude at the Rx PIPED and the complex T-wave current amplitude at the Tx PIPED. The magnitude of the complex transfer impedance is directly proportional to the baseband current amplitude $I_{BB,1}$

$$|\underline{Z}_{21}| = a\hat{I}_{BB,1}, \tag{55}$$

where the proportionality factor a depends on the operation conditions of the Tx and Rx PIPED (Supplementary Section 7 in [6]):

$$a = \left(\frac{1}{2} \frac{dS_{Rx}}{dU_{Rx}} \bigg|_{U_{Rx,0}} \hat{P}_{Rx,1} S_{Tx}\left(U_{Tx,0}\right) \hat{P}_{Tx,1} \right)^{-1}. \tag{56}$$

The corresponding phase φ_{21} is obtained from φ_{BB} by swapping the role of the Tx and Rx PIPEDs, thus allowing us to eliminate the influence of the unknown group delay in the optical fibers (see Methods in [6]). The measured magnitude and phase characteristics of the transfer impedance $\underline{Z}_{21,meas}$ are depicted as blue traces Figure 23 (c-e). For comparison, we also numerically calculated the transfer impedance of the T-wave transmission lines using a

commercially available time-domain solver. The results of the simulated transfer impedance
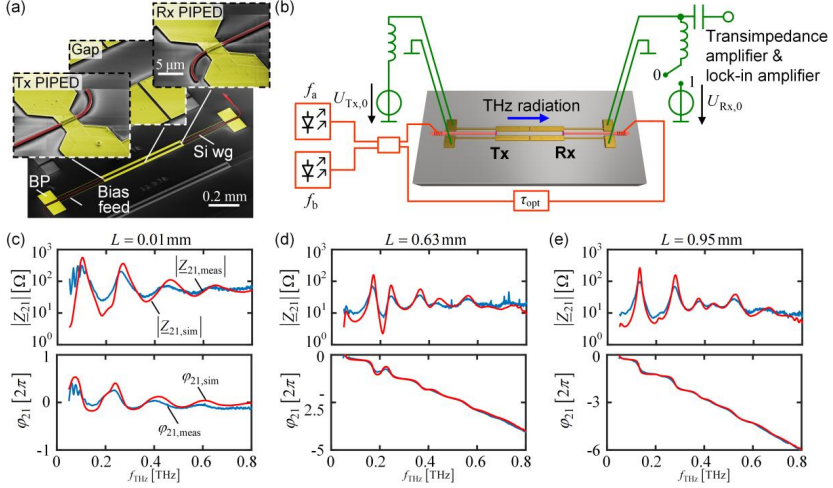

Figure 23. Demonstration of T-wave system on a silicon chip. **(a)** False-color SEM image of the terahertz system, consisting of a T-wave Tx, a coplanar transmission line, and a T-wave Rx. Light is coupled to the PIPED Tx and Rx by silicon waveguides (Si wg) and grating couplers (GC). Bias pads (BP) and terahertz-blocking bias feeds (terahertz chokes) allow a voltage to be applied to or a signal current to be extracted from the PIPED. A gap in the middle of the transmission line separates the bias voltages of Tx and Rx. Insets: magnified SEM images of Tx, Rx and the transmission line gap. **(b)** Experimental set-up for on-chip transmission-line measurements. Tx and Rx are fed with two optical spectral lines f_a and f_b at varying frequency separations. The Tx is biased in the forward direction $U_{Tx,0} > 0$ and feeds the T-wave transmission line. At the end of the transmission line, the Rx PIPED acts as a homodyne receiver to coherently detect the T-wave signal by down-conversion to the baseband current $I_{BB,1}$. To reduce the noise, the switch is set to 0 to operate the Rx PIPED with zero-bias current. **(c)–(e)** Magnitude and phase characteristics of the complex transfer impedance $\underline{Z}_{21} = |\underline{Z}_{21}| \exp(j\varphi_{21})$. Red curves show simulated frequency characteristics of $|\underline{Z}_{21,sim}|$ and $\varphi_{21,sim}$ for three transmission line lengths $L = 0.01$, 0.63 and 0.95 mm, and blue curves represent corresponding measurement results $|\underline{Z}_{21,meas}|$ and $\varphi_{21,meas}$. The frequency roll-off of the PIPED is modelled by an RC low-pass filter with a 3 dB frequency of $f_{3dB} = 0.3$ THz (Appendix of [6]). Measurements (blue) agree very well with the simulation (red).

$\underline{Z}_{21,\mathrm{sim}}$ are shown as red traces in Figure 23 (c-e). Note that calculating the proportionality factor a according to Equation (56) is subject to large uncertainties because the experimental operating conditions of the Tx and Rx PIPED at terahertz frequencies are only known approximately. To estimate $|\underline{Z}_{21}|$ from the measured baseband photocurrent amplitude $I_{\mathrm{BB,1}}$, we therefore chose a such that we obtained the best agreement between the simulation and measurement (Supplementary Section 8 in [6]). Moreover, we numerically corrected for the frequency roll-off of the PIPED when deriving the measured transfer impedance $\underline{Z}_{21,\mathrm{meas}}$ from the baseband photocurrent amplitude $I_{\mathrm{BB,1}}$. To this end, the transit-time limited frequency response of both PIPEDs was approximated by an RC low-pass characteristic with a corner frequency of 0.3 THz, obtained from a least-squares fit of the measurement data (Supplementary Section 8 in [6]).

The measured and simulated transmission line characteristics show perfect agreement with respect to amplitude and phase over the entire frequency range $f_{\mathrm{THz}} = 0.05 - 0.8\,\mathrm{THz}$, corresponding to a span of 1.5 decades. The field distributions along the transmission lines were investigated by numerical simulations of the structures, see Supplementary Section 10 in [6]. These findings show that signal processing in PIPED can be quantitatively described by a reliable mathematical model, thus enabling deterministic photonic–electronic signal processing over an ultra-broadband frequency range.

6.6   Summary and Outlook

We have demonstrated a novel approach to T-wave signal processing that exploits internal photoemission at metal–semiconductor interfaces of silicon–plasmonic structures. This approach allows to monolithically co-integrate T-wave transmitters and receivers on the silicon photonic platform, thus enabling a novel class of photonic–electronic signal processors that may exploit the outstanding technical maturity and performance of the silicon photonic device portfolio. When used as a T-wave Rx, the performance of our current PIPED devices can already compete with state-of-the-art III–V photoconductors, while further improvements are still possible by optimized device geometry and materials. The devices are analyzed and described by a quantitatively reliable mathematical model. In a proof-of-concept experiment,

we use co-integrated PIPED transmitters and receivers to measure the complex transfer function of an integrated T-wave transmission line.

[*end of paper*]

7   Summary and Outlook

This thesis introduces and experimentally demonstrates a novel approach of exploiting internal photoemission (IPE) for ultra-broadband photodetectors. For brevity, we name the resulting device a PIPED, an acronym for **Plasmonic Internal Photoemission Detector**. PIPED represents the first detector based on IPE that operates at technically viable performance figures. The excelling performance is due to its specific structure that combines a metal-semiconductor-metal (MSM) with a metal-oxide-semiconductor (MOS) junction similar to a MOS-field-effect transistor. This combined MSM/MOS junction has nano-scale dimensions and allows PIPED to combine surface plasmon polaritons (SPP) and internal photoemission in a virtually ideal way: Light is coupled from a silicon photonic waveguide and transformed to a surface plasmon polariton in a metal-semiconductor-metal junction made of gold, silicon and titanium. The silicon core in this layer stack is less than 100 nm wide. Gold and titanium have vastly different absorption characteristics, which leads to a strong SPP absorption at the Ti surface, depositing most of the incident optical power there. Due to the strong localization of SPP, hot electron-hole pairs are created in the metal directly at the interfaces - close enough to allow a hot carrier transfer across the interfaces, thus creating a photocurrent. Applying voltages across the MSM or the MOS junction allows to control the electronic band structure and thus the internal carrier distribution, further improving the device performance. More specifically, applying a voltage across the MSM junction increases the hot carrier transfer probability and enables carrier tunneling through the barrier, while a voltage across the MOS junction creates a highly conductive carrier channel and increases the optoelectronic bandwidth to the terahertz frequency range (T-waves).

The extremely high bandwidth of PIPED enables the generation of T-waves by photomixing, exploiting the beating of two optical carrier tones with a frequency difference in the THz domain. PIPED can also act as a coherent T-wave receiver: The pronounced voltage dependency of the sensitivity allows to down-convert T-waves by superposition with an optically created local

oscillator. Hence, PIPED are viable candidates for optoelectronic T-wave systems.

The key accomplishments of this work are summarized below, beginning from device aspects, and moving to system aspects:

First demonstration of a silicon plasmonic photodetector with technically viable performance figures: We experimentally demonstrated PIPED with a record-high sensitivity of more than 0.12 A/W at a wavelength of 1550 nm and with large optoelectronic bandwidths well above 40 GHz. We expect that the sensitivity can be increased to at least 0.24 A/W by controlling the input polarization of the light. We prove the viability of the PIPED concept for real-world applications by receiving on-off keying (OOK) data at a rate of 40 Gbit/s. This performance was orders of magnitude above the state of the art for this class of devices at the time of its publication.

Development and implementation of a straightforward PIPED nano-fabrication scheme in silicon: Our fabrication process is able to create PIPED with a device footprint of below $1\,\mu m^2$. The device remains stable for silicon core widths even below 20 nm, which matches the capabilities of latest semiconductor fabrication techniques [92]. This makes PIPED much smaller than competing Si/Ge or UTC photodetectors, which typically consume space in the order of $100\,\mu m^2$ [19]. Hence, PIPED allow for an unprecedented optoelectronic integration density.

Ultra-wide range of operation: The PIPED operation range is only limited by the absorption band edge of silicon and the barrier height of the metals. For our PIPED design, we expect a useful operation range from 1130 nm to more than 2000 nm. We have yet to determine the optimal operation wavelength that maximizes the sensitivity.

Ultra-high optoelectronic bandwidth beyond 1 THz: We show that a standard transit-time limited PIPED already features a bandwidth of the order of 300 GHz. This bandwidth can be significantly increased: A gate electrode allows to induce a conductive channel in the device where its dynamic response is determined by dielectric relaxation instead of the carrier transit time. We show that this concept is effective even at T-wave frequencies. Our simulations

show that by using this gate electrode, the optoelectronic bandwidth can be increased to well above 1 THz.

First demonstration of silicon plasmonic T-wave transmitters and receivers: Exploiting the high bandwidth and the voltage-dependent sensitivity, the same PIPED device can generate T-waves by photomixing and receive T-waves by optoelectronic down-conversion using an optically generated local oscillator. Using this technique, our PIPED is capable of receiving T-waves with frequencies above 1 THz, even when no additional field-effect is employed.

First demonstration of integrated T-wave systems: We show that PIPED can be monolithically cointegrated with T-wave circuits. As a proof-of-concept experiment, we create a T-wave system comprising two PIPED that are connected by an on-chip T-wave transmission line. One PIPED is operated as a T-wave transmitter while the other acts as a T-wave receiver. We show that this setup allows to measure the complex-valued transfer characteristics of the transmission line in the frequency range from 0.05 THz to 0.8 THz.

7.1 Outlook

The achievements of this work are enabled by the unique characteristics that PIPED offers for silicon-plasmonic photodetection. However, the developed devices represent only our first-generation proof of concept. Conceptual adaptions that seek to improve the performance and that address the device's limitations are possible. In the following, we list a few selected starting points for future developments.

CMOS compatibility: The fabrication process that we employed in this work is fundamentally not CMOS compatible, mainly due to the presence of gold. The presence of gold is not permissible in silicon foundries due to its strong propensity to diffuse into silicon. The metal evaporation and deposition at grazing incidence is also not a standard technique in the CMOS tool set. However, the fabrication process can be made CMOS compatible in principle. The gold drain electrode could be replaced by copper, which is a standard material in CMOS processing and features similar plasmonic characteristics. Other options for the titanium source electrode could be metal silicides like

titanium silicide or tungsten silicide. These materials are typically used to create an efficient ohmic contact between metals and semiconductors by reducing the potential barrier at the interface. This is a feature of particular interest when it comes to internal photoemission.

Output current, damage threshold, and thermal coupling: The maximum current that we could extract from the PIPED in our experiments was limited by a relatively low damage threshold for an individual device. The small footprint of PIPED leads to very high current densities, even at moderate optical input powers. Our post-failure analyses of damaged devices revealed that the most prominent failure mode was melting of the source and drain electrodes, indicating extreme temperature conditions at the time of device breakdown. This issue can be addressed by several measures. First, a layer of a surrounding material with a high heat capacity and conductivity could act as a heat sink, reducing the temperature stress imposed on PIPED devices during operation. Second, also device-internal measures might help. Heat is a consequence of inelastic collisions of carriers, for example due to carrier scattering at rough interfaces. Improving the sidewall roughness of the PIPED interfaces might already significantly decrease the power dissipation in the device. A more advanced approach might be found in the deposition of two-dimensional layers on the silicon core. The reduced density of electron states in the two-dimensional layer reduces the electron-phonon scattering [93]. Hence, usage of such layers might allow to reduce the interaction of the electronic system and the phonon system, the excitation of which is the strongest contributor to a device heating.

Another possible approach to increase the output current is to distribute the optical input power among an array of PIPED, where the individual source (or drain) electrodes are interconnected, e.g., in an interdigitated structure. The overall receiver current is then the sum of all individual PIPED current contributions. State-of-the-art semiconductor processing is capable to create dense layer arrays of MSM junctions back-to-back. Hence, the total device footprint would still be small.

High-κ dielectrics as gate oxide: We used SiO_2 as a gate oxide, mainly due to its excellent properties for device fabrication. However, the permittivity of SiO_2 is relatively low. Using high-κ dielectrics like Al_2O_3 as gate oxide instead might

allow to significantly reduce the required voltage levels for operation on the gate. Such reduction is beneficial for future operation schemes where a weak T-wave directly controls the gate voltage. In such an application, a T-wave fed to the gate contact can act as a local oscillator in the silicon plasmonic coherent receiver. High-κ dielectrics can also be used to improve the optical properties of the device, for example by increasing the field enhancement in the oxide layer. A higher field in the oxide improves the field localization at the metal apexes and hence can lead to a further increased sensitivity.

We believe that future PIPED with CMOS-compatible fabrication, with an improved damage threshold, and with improved materials will open a path to densely integrated optoelectronic systems operating at unprecedented efficiency and bandwidth.

Appendices

A. Fabrication and Characterization Technology for High-Speed Plasmonic Photodetectors

This appendix provides supplementary information (SI) to the general tools employed for the fabrication and characterization of silicon plasmonic photodetectors. The content is largely identical to the SI of our original paper [J6] in which we reported about PIPED:

Silicon-Plasmonic Internal-Photoemission Detector for 40 Gbit/s Data Reception – Supplementary Information

S. Muehlbrandt[1,2,*], A. Melikyan[1,2], T. Harter[1,2], K. Köhnle[1,2], A. Muslija[1], P. Vincze[3], S. Wolf[2], P. Jakobs[1], Y. Fedoryshyn[4], W. Freude[2], J. Leuthold[4], C. Koos[1,2,*], M. Kohl[1,*]

We have used this SI as a basis, and we supplemented further content that we developed in the meantime.
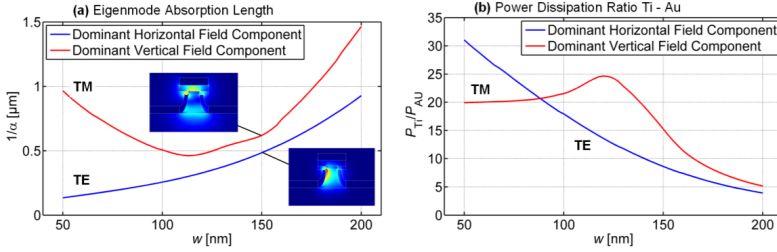
Figure 24. PIPED eigenmode attenuation characteristics. **(a)** Absorption length in the detector core as a function of the core width for the eigenmodes with dominant horizontal (TE) and vertical field orientation (TM). The insets exemplarily show the associated optical field magnitudes at a silicon core width of $w = 150$ nm. **(b)** Power dissipation ratio of Ti over Au. The fraction of light absorbed in the Ti is significantly larger than the fraction in the Au.

A.1 Power Attenuation Constant and Device Length

The power attenuation constant α in the photodetector core determines the lower limit for the minimum technically useful device length. We use a two-dimensional finite-element method to numerically compute the complex-valued effective refractive index $\underline{n}_{\text{eff}} = n_r - \mathrm{j}n_i$ associated to the eigenmodes ψ of the MSM waveguide with dominant horizontal and dominant vertical field orientation. We use experimentally obtained data for the complex frequency dependent permittivities of Au [17] and Ti [18]. The details for the numerical modelling can be found in Appendix C.

The imaginary part of the effective refractive index n_i and the free-space wavenumber $k_0 = 2\pi/\lambda$ define the power attenuation constant $\alpha = 2k_0 n_i$. The associated absorption length for which the optical intensity has decayed by a factor of $1/e$ is given by $l_{\text{abs}} = 1/\alpha = 1/(2k_0 n_i)$.

Figure 24 (a) shows the absorption lengths for the eigenmodes of the detector core with dominant vertical or horizontal field alignment as a function of the detector core width w. The absorption length is significantly smaller than 1 µm for $w < 150$ nm and for either field orientation. This is due to an ever-increasing field confinement inside the detector core as the junction width w is decreased.

A large field overlap with the absorbing Ti layer leads to a rapid absorption of light. The absorption length increases again for the dominant vertical field orientation at widths $w < 100$ nm as the mode is increasingly guided outside the MSM junction, farther away from the Ti. Taking five absorption lengths as a basis, detectors with a length between (2.5 ... 5) µm absorb more than 99 % of the optical field, depending on the detector width and the polarization. These short absorption lengths enable ultra-compact plasmonic photodetectors.

The analyses from the Sections 4.4 and 5.5.1 indicate that Ti is a good photoemitter, especially under high biasing conditions. Hence, a high localization of light absorption in the Ti is crucial for a high performance. This condition is naturally met in the PIPED design. Figure 24 (b) shows the ratio of optical power absorbed in the Ti and the power absorbed in the Au. The power absorption has been evaluated from the eigenmode and the field overlap with the metals, see Appendix C.1 for details. Ti absorbs significantly more power throughout the entire range of practical core widths w. For the TM mode, the highest ratio reaches a value of $25\times$ at $w = 125$ nm. The decrease for even smaller $w < 125$ nm corresponds to the decreasing absorption length in Figure 24 (a), which might be due to a smaller field overlap with the drain and source electrodes for the TM mode and for very small w. However, the power dissipation ratio of the TE mode steadily increases for smaller w, even exceeding a value of $30\times$ at $w = 50$ nm. We analyze the distinct power distribution in the full PIPED structure in the following section.

A.2   Optical Coupling and Absorption

A suitable distribution of light absorption is a deciding factor for the optoelectronic conversion performance of the PIPED. The ideal detector geometry deposits all optical power at the titanium, which is a better photoemitter than gold, and at the top of the detector core, where the separation of the metals is the smallest. Conversely, the light absorption in the tapered mode converter should be minimized, as the carrier extraction efficiency is smaller due to the metal separation being larger than in the core. We analyze the absorption characteristics of our device by a numerical simulation. The simulation uses a full-vectorial 3-dimensional simulation of the PIPED, including the silicon
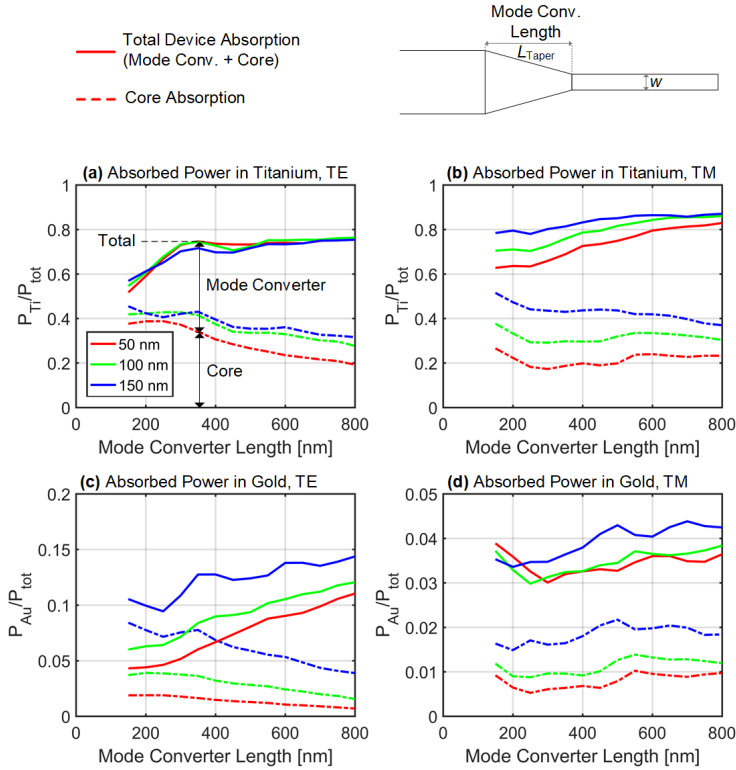
Figure 25. PIPED simulated absorption characteristics for different mode converter lengths and core widths *w* in the range from 50 nm (red curves), 100 nm (green curves) and 150 nm (blue curves). The analysis distinguishes the total power absorbed in the detector (full lines) and the detector core (dashed lines). The difference between the two associated lines is the power absorbed in the mode converter. The absorbed power fractions are further distinguished between the metals titanium (subfigures **(a)** and **(b)**) and the gold (subfigures **(c)** and **(d)**), as well as for the polarization (TE in (a) and (c), TM in (b) and (d)). All subfigures show that the power absorption in the Ti dominates for the entire parameter range under analysis.

photonic input waveguides, the tapered mode converter, and the detector core. The details of the numerical modelling environment can be found in Appendix C.

The model computes the optical power dissipation in each metallic compartment of the PIPED, i.e., the power deposited in the titanium and the gold, and distinguishes the power absorbed in the mode converter and the core. We employ the mode converter length L_{Taper} and the detector core width w as simulation parameters and select them in the typical range from $L_{\text{Taper}} = 150\,\text{nm} \dots 800\,\text{nm}$, and $w = 50\,\text{nm} \dots 150\,\text{nm}$. The simulation is executed for the TE and the TM polarization.

The results can be found in Figure 25, subfigures (a) through (d). The Ti dominates the absorption for the entire range of simulated parameters and for all polarizations, featuring an order of magnitude larger power absorption than the Au. Hence, we will focus on the absorption in the Ti in subfigures (a) and (b) in the following. For both the TE and the TM polarization, very short mode converters result in a relatively low total absorption in the device, indicating large reflection at the input of the detector. This is not unexpected due to the large difference of the effective indices in the silicon photonic input waveguide and the plasmonic one, as well as the vastly different field distributions.

The total absorbed power in the device (full lines) increases with increasing mode converter length. This indicates an optimized coupling to the device due to a more efficient mode conversion from the silicon photonic waveguide to the plasmonic section. However, the absorbed power in the core decreases with increasing mode converter length. This is expected, as the mode converter in itself is strongly absorbing and dissipates a large fraction of the incident optical power. The longer the mode converter is, the less power can reach the core (dashed lines).

Specifically for the TE polarization, the absorption characteristics do not vary strongly across different core widths. It can be seen that the total absorbed power does not increase anymore when the mode converter is longer than $L_{\text{Taper}} = 350\,\text{nm}$, while the power absorbed in the core decreases.

Surprisingly, shorter mode converter lengths $L_{\text{Taper}} < 350\,\text{nm}$ decrease the total amount of absorption in the device, but increase the fraction of absorbed power in the device core. Following this finding would mean that a very short or no mode converter at all should work best in terms of optoelectronic conversion. However, this doesn't match the experimental results we obtained for the

variation of the mode converter lengths, indicating no distinctively improved performance for PIPED with short mode converters.

We believe that the reason for this apparent contradiction might be that above power dissipation analysis does only account for the absolute power deposited in the metal, integrated across the entire metallic domain. However, only the fraction of power absorbed in close proximity of the metal apex has relevance for the optoelectronic conversion. This understanding is in line with our results for the optoelectronic conversion efficiencies of the same PIPED operated in TE or TM polarization in Section 4.5 *Optical Coupling and Photocurrent Polarization Dependence*. The photocurrent for the TM mode was nearly twice the as large as for the TE mode, due to the better excitation of photocarriers close to the metal apexes. Furthermore, our simulation shows that the PIPED with no converter or a short converter in fact comprises significant electric field strength in the core, but it is located far away from the apex. It is expected that light absorbed in the core but away from the apex does not contribute substantially to the photocurrent.

For the TM polarization, Figure 25 (b), there is only a weak dependency of power absorbed in the core and the mode converter length. The curves for a given core width *w* are relatively flat. However, there is a more pronounced dependency on the width itself. It seems that it is increasingly difficult to couple power to ever smaller cores.

Note that the results discussed in this section are related to a similar discussion in Section 4.5, Table 7. Compared to our previous discussion, the results in this section have been created from an updated simulation environment, where the definition of the PIPED structural features have been improved. Consequently, this leads to slightly different numerical results with no substantial impact on the main findings. (Example: Total absorption in the Ti P_{Ti}/P_{Tot} for TE polarization and $L_{Taper} = 550$ nm according to the update: 0.73, previous results 0.80.)

We conclude that the method presented above is helpful to understand the principles of coupling to the PIPED device by means of an adiabatic mode converter. However, the modelling does not allow to establish an immediate relation between the absorbed power and the resulting opto-electronic

conversion efficiency. To achieve this, the model must be supplemented to relate the local light absorption to the carrier dynamics in the metals, including the emission efficiency across the MS-barrier. As of now, such comprehensive model does not exist to the best of our knowledge.

A.3 Fabrication

The fabrication of ultra-narrow MSM waveguide structures with semiconductor layer thicknesses below 100 nm, smooth sidewalls and different metal layers without direct contact poses a considerable challenge on nanofabrication. While direct etching of metal-coated waveguides via a reactive-ion etching (RIE) or
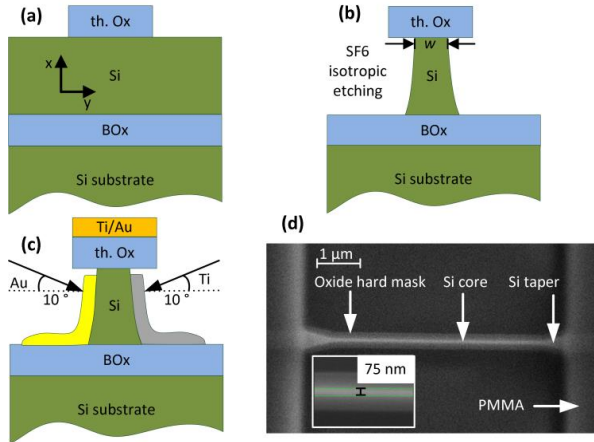
Figure 26. Fabrication sequence of plasmonic internal-photoemission detectors (PIPED) **(a)** 80 nm thick thermal oxide is grown and patterned on top of a weakly *p*-doped silicon on insulator (SOI) wafer with a device layer height of 340 nm in x-direction. The Si height after oxidation is 300 nm. **(b)** The hard mask is undercut using isotropic etching of silicon with an SF6 chemistry at room temperature. The hard mask is unaffected due to the high selectivity of SiO_2 over Si. The ratio of vertical to lateral etch rate is ~3. The Si core remains mechanically stable for the width w = 75 nm. **(c)** Au or Ti are deposited on the sidewall of the Si core under grazing incidence. The hard mask serves as an evaporation shadow preventing electrical shortcut. **(d)** SEM image of the undercut hard mask before metallization. The hard mask is translucent and reveals the underlying Si core. The inset shows a magnified image of the silicon core with a width of w = 75 nm. The Si tapers connect the plasmonic section to photon waveguides on either side of the detector. The photonic parts are protected from metallization by a PMMA resist layer.

reactive ion beam etching (RIBE) process suffers from poor selectivity of resist over metal, multiple lift-off processes become very critical in the sub-100 nm range. Therefore, we introduce a new process that is based on direct metal evaporation at self-masking silicon nanowires.

Here, the process is demonstrated for silicon-on-insulator (SOI) chips with a low p-doped 340 nm silicon device layer (resistivity ~10 Ωcm). The SOI chips are first thermally oxidized to create an 80 nm thick top oxide layer. The oxide layer is patterned using electron beam lithography and RIE with a CHF$_3$ chemistry, generating a hard mask for the layout of photonic and plasmonic sections on the residual 300 nm of the silicon device layer (Figure 26 (a)). Subsequent electron beam lithography with PMMA resist opens the plasmonic sections and defines the contour of the electrodes for external electronics. The resist layer serves as a mask for an isotropic etching step with SF$_6$ at room temperature (Figure 26 (b)). The SF$_6$ only attacks the silicon in the opened plasmonic and electrode sections and leaves the oxide unharmed. The oxide hard mask is undercut and the electrode areas are etched down to the buried oxide. As a result, a silicon nanowire with an oxide cap is created. Due to the isotropic nature of this etching step, the surface roughness of the silicon is well below the roughness of the original mask of below 20 nm. The structure is mechanically stable for lateral nanowire widths of $w = 75$ nm and below. In a further step, the undercut oxide serves as a hard mask for metal evaporation at grazing incidence with respect to the plane of the chip from both sides of the nanowire using different metals (Figure 26 (c)). Under such condition, an asymmetric MSM junction is created without an electrical short circuit. The metal thickness of approximately 40 nm was chosen to exceed the SPPs penetration depth. With a carefully designed PMMA mask for angled metallization there is no fundamental limit in the device length, and devices with a length $L < 5$ µm become possible. This processing is not restricted to metals. In principle, any desired lateral layer sequence can be achieved, given that the materials can be directionally deposited. After metal lift-off, the remaining silicon is fully etched with an anisotropic SF$_6$ etching step at cryogenic temperatures. The previously patterned oxide hard mask is used to transfer the photonic waveguides und the diffraction grating couplers into the silicon device layer. The photonic waveguides have a width of 400 nm and a Si
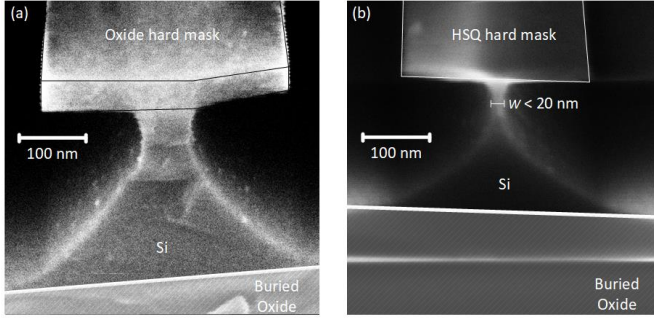
Figure 27. Comparison of PIPED with different hard mask materials. **(a)** Oxide hard mask prepared from thermal oxidation of silicon. The 80 nm thick oxide layer provides excellent masking properties. It remains stable even for undercuts that leave only $w = 75$ nm of Si. Undercuts with a remaining silicon width of less than 50 nm make the structure unstable. This is due to the thermal stress that is induced during the oxidation process. **(b)** Hard mask structured from an HSQ resist layer. The HSQ mask is not as resilient as the oxide and the layer height is depleted during the Si etch. However, as the HSQ is structured without tensile stress, the mask remains stable even for very small Si widths w > 20nm. The ratio of Si feature width and lateral mask width is larger than 10:1.

height of 300 nm. The 80 nm thick oxide hard mask is not removed. Figure 26 (d) shows an SEM image of a plasmonic section before metallization.

The plasmonic sections are connected to Si photonic waveguides by mode converters that are modelled as MSM adiabatic tapers. This processing scheme allows for fabrication of lambda-scale active plasmonic devices with electrode distances below 75 nm.

The minimum lateral width w of the Si is limited by the tensile stress of the oxide hard mask. The oxide layer is created at high temperatures in excess of 1000 °C by thermally oxidizing the silicon layer. The Si and the SiO_2 layer accumulate a tensile stress during the cool down phase after the oxidation. The adhesion between the layers is strong enough to prevent delamination. However, when the oxide layer is undercut, the interface between the strained SiO_2 and the Si becomes smaller and the capability of the Si to hold the SiO_2 in place decreases. We have found that in the structural configuration used in our experiments (oxide layer thickness 80 nm, oxide layer width 340 nm) the SiO_2 does typically not delaminate at remaining silicon widths of $w = 75$ nm, see Figure 27 (a). The lateral aspect ratio for such structure is 4.5:1 (width of hard

mask:width of Si). The probability of delamination increases when w decreases and structures with $w < 50$ nm become gradually more unstable during the following process steps.

PIPED with smaller widths $w < 50$ nm become feasible by eliminating the tensile stress between the hard mask and the Si layer. This can be achieved by replacing the thermal oxidation process step by deposition of a hard mask. A suitable replacement for SiO_2 is HSQ, which is a negative tone resist. The HSQ layer is spun on the Si. The hard mask preparation is completed by electron beam lithography and resist development. We found in our sample preparation campaigns that PIPED prepared from HSQ allow for Si widths even below $w < 20$ nm, see Figure 27 (b), while leaving the other structural parameters width and thickness of the hard mask unchanged. This corresponds to a lateral aspect ratio of more than 17:1. This means a 340 nm wide layer of HSQ is suspended by a silicon pedestal of less than 40 atomic layers measured at its narrowest diameter.

We found that the HSQ utilized in our campaigns did not provide the same excellent properties as compared to the thermally oxidized SiO_2 hard mask. Figure 27 (b) shows that the layer thickness of the HSQ is significantly decreased from the initial 80 nm thick layer after the etching, indicating a worse etch resilience of HSQ compared to SiO_2. Furthermore, our opto-electronic characterization of PIPED prepared with HSQ hard masks showed that neither the sensitivity, nor the resilience against dielectric breakdown amount to the performance of PIPED prepared with SiO_2 hard masks. To replace the SiO_2 hard mask, our processing techniques for HSQ must first be improved. Alternatively, other materials with additional function, like high-κ-dielectric materials, might even turn out a better pick in regards of the PIPED performance. They might be able to not only provide an etch-resilient hard mask, but could improve the field-effect characteristic of FE-PIPED (see Chapter 5 "*Field-Effect Silicon Plasmonic Photodetector for Coherent T-wave Reception*").

A.4   Opto-Electronic Characterization

The following sections explain how we characterize our devices after fabrication. Light at the wavelength of 1550 nm is amplified in an Er-doped fiber amplifier (EDFA) and coupled into the fabricated chip by diffraction grating couplers. The photonic waveguide mode is excited with a dominantly horizontal electric field component. The coupling losses of the grating couplers and the propagation losses of the photonic waveguides are assessed with the subtraction method. The transmission T of photonic reference waveguides is determined as a function of waveguide length l, see Figure 28. The transmission as a function of waveguide length is a linear function on a semi-logarithmic scale. The slope m of this function is the waveguide propagation loss amounting to $m = 2$ dB/mm. The y-intercept T_0 corresponds to two times of the grating insertion loss, $T_0/2 = 10.1$ dB. The high propagation losses of the SOI waveguides are due to the surface roughness of the non-optimum negative-tone resist that was used to structure the SiO2 hard mask. This loss contribution can be reduced to the order of 1 dB/cm using high-resolution resists like HSQ [94].

We use an RF probe with ground signal (GS) electrode configuration to pick up dark current I_d and photocurrent I_p from the device. The photodetector is biased with a precision source/measure unit (Agilent B2960A). For DC measurements, the current-voltage characteristic is measured under
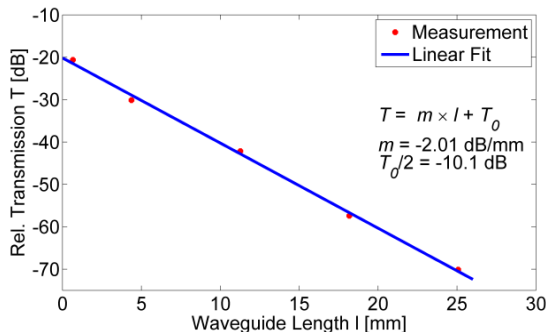

Figure 28.   Measurement of photonic waveguide propagation loss and grating coupler insertion loss.

illumination at various power levels P. Measurements at dark conditions serve as a reference to assess the photocurrent $I_p = I - I_d$ and the sensitivity $S = I_p / P$.

For Polarization sensitive measurements, we cleave SOI chips and access the photonic waveguides from the chip edges using polarization maintaining (PM) fibers. The slow polarization axis of the fiber is aligned with the surface normal of the photonic chip. For alignment of the coupling fiber, we use a rotatable fiber holder and a free-space polarizer. The orientation of the linear polarization at the input of the PM coupling fiber is then controlled with a fiber-coupled half-wave plate (Thorlabs PC FFB 1550). The half-wave plate allows rotating the polarization corresponding to a full 360° revolution of the electric field vector at the input of the PM fiber. Due to the strong birefringence of the PM fiber, only input electric fields parallel to one of the polarization axes represent a linear polarization at the input of the photodetector.

For opto-electronic bandwidth measurements, a bias tee is inserted into the electronic circuitry after the probe to split electronic DC and RF signals. A 2-port vector network analyzer (VNA, Anritsu 37397C) is used to drive a Mach-Zehnder Modulator (MZM), creating an intensity modulated signal. The optical signal is received by the PIPED and the electrical signal is returned to the VNA. The RF-network is referenced with a photodetector standard (Anritsu MN4765A). The RF-probe is not referenced for.

A.5   Data Reception Experiment

Infrared light at the wavelength of 1550 nm is launched into a MZM. The intensity of the optical signal is encoded with a pseudo-random bit sequence (PRBS) with a bandwidth of up to 40 Gbit/s, resulting in an on-off keying optical signal. This is the highest data rate directly available in our labs. The modulated light is amplified by an EDFA, then coupled to the Si chip and received by the biased PIPED. The electric signal is directly amplified after the RF-probe and a sampling head of a digital communication analyzer records the eye diagram (Agilent 86100 Infiniium DCA-J C).

B.   PIPED Design Validation

The design of PIPED specifically aims at exploiting internal photoemission for high-speed photo detection. While the design is the key to the excelling performance of PIPED, the previous chapters have rather focused on its technological benefits as compared to the validation of PIPED design principles. However, investigating the physical conditions of internal photoemission in the device is essential to establish device models and to improve the performance of succeeding PIPED device designs.

In this appendix, the photocurrent dependency on wavelength and optical input power of metal-semiconductor-metal (MSM) junctions is analyzed. To this end, different MSM test designs are created and their photocurrent response is recorded. From the results it can be inferred that in PIPED devices, most of the photocurrent stems from single hot carriers instead of a heated carrier gas. We furthermore find that PIPED operates across an optical wavelength range from $\lambda = 1300$ nm to at least 1600 nm, and we expect the usable range to cover an even larger wavelength range from the Si absorption-band edge at 1130 nm to beyond 2000 nm.

Premise

From a phenomenological perspective, light enters the PIPED device and creates a measurable photocurrent in the external circuit. The photocurrent is created due to photon absorption at metal-semiconductor interfaces and subsequent carrier injection from the metals into the semiconductor. However, it is yet unaddressed to which extent the emission is due to single hot carriers or due to a hot carrier gas at an elevated temperature T_c.

The understanding of the internal sources of the photocurrent sets guidelines for future device designs: If single hot carriers constitute the majority of the photocurrent, then device designs that increase the lifetime of hot carriers could increase the carrier extraction efficiency since the single carriers would have more time to attempt the crossing of the MS-interface. Such increase in lifetime could be achieved by controlling the local density of states, e.g., by shaping the

metal electrode into two-dimensional patterns. The reduced density of electron states in the two-dimensional layer reduces electron-electron, as well as electron-phonon scattering, which increases the carrier lifetime [93].

Conversely, if thermionic emission of a hot carrier gas generates most of the photocurrent, then the extraction efficiency might be improved by insulating the hot-carrier gas from thermalizing with the surrounding phonon system. The following sections explain the details of our analysis.

B.1   Methodological Approach and Setup

We use Equation (23) (Fowler photoemission) and (21) (thermionic emission) to discriminate the relative current contributions of single hot carrier emission and emission from a heated carrier gas. Both equations predict a distinctively different photocurrent, rendering the effects distinguishable in the experimental data:

- If single hot carriers dominate the photocurrent, then Equation (23) holds. In this case, the photocurrent should increase exponentially with the photon energy (or inverse exponentially with wavelength) and linearly with the power at a given photon energy.
- If the heated carrier gas dominates the photocurrent, then Equation (21) holds and the photocurrent would increase with the carrier gas temperature according to a square-law. It is fair to assume that the temperature of the carrier ensemble increases linearly with the total amount of deposited power. In this case, it does not matter whether the additional power is provided by an increase of photon energy for a given number of photons, or from an increase of the number of photons at a given photon energy.

The data presented in this chapter are based on three different MSM device designs *I*, *II*, and *III*, which represent intermediate prototype-stages of the PIPED device development. These early devices were already able to output photocurrents under light illumination and hence could be used to characterize the MSM interfaces. However, the performance was not yet up to the record results obtained from the final PIPED and FE-PIPED designs reported about in
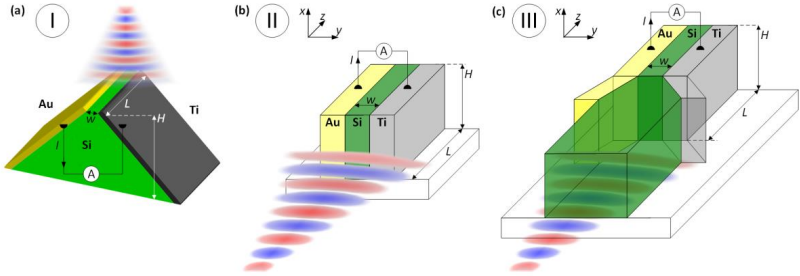
Figure 29. MSM designs employed for test structures. The designs represent intermediate stages from the development before the final PIPED design **(a)** Wedge structure. The MSM junction is created on cleaved shards from Si wafers. The MS interfaces are so large that dark currents would dominate under applied voltages. **(b)** PIPED without feed waveguide. The devices are obtained from fabrication of long MSM junctions and cleaving, thereby exposing the facet. The device lengths are in the order of 5 mm, hence voltages cannot be applied due to excessive dark currents. **(c)** PIPED including feed-waveguide. The device lengths are short enough to allow applying voltages. However, the device facet is not accessible and coupling is achieved via the feed waveguide. The oxide cap and the gate electrode have been omitted in (b) and (c) for better visibility.

Table 8. Properties of PIPED designs.

Design	I, Figure 29 (a)	II, Figure 29 (b)	III Figure 29 (c)
w	~50 nm – 150 nm	75 nm – 200 nm	75 nm – 200 nm
L	~500 µm	~5 mm	~ 10 µm
H	10 mm	300 nm	300 nm
Light coupling	Via exposed PIPED facet	Via exposed PIPED facet	Via waveguide facet
Voltage applicable?	No, dark current dominates.	No, dark current dominates	Yes

the Chapters 4, 5 and 6. The structure and properties of the device designs are summarized in Figure 29 and Table 8.

We use a cleaved fiber to supply the light and assume that the light spot is always much larger than the facet of the MSM junction and that the fraction of light coupled to the MSM junction for various wavelengths and for a given

power in the fiber is essentially the same. With this assumption, we neglect a wavelength-dependent coupling loss to the MSM junction.

We use a wavelength-tunable laser source that spans the range from 1270 nm – 1420 nm, corresponding to photon energies *hf* in the range from 0.868 eV – 0.977 eV. The lower wavelength (higher photon energy) threshold is sufficiently far away from the Si-band edge and direct inter-band absorption of Si can be excluded. The higher wavelength threshold is limited by the performance of the laser source. We use a second laser source to couple light at 1550 nm.

We expect a strong polarization-sensitivity of the photocurrent for all tested MSM designs due to the plasmonic properties of the detectors. We manually optimize the polarization for highest photocurrent for measurements with test Design I. For measurements with the Designs *II* and *III*, we use a polarization-maintaining optical setup and a free-space polarizer to select the input polarization. We do not apply an external voltage to the Designs *I* and *II*. These designs are much longer as Design III, and an external voltage would lead to an excessive dark current. Without a bias voltage, we only expect electron emission from Au into Si according to the evaluation of Chapter 2.2 The device current is recorded using a source-measure-unit (SMU).

B.2   Experimental Results

Wavelength and Power Dependence, Sample Design *I*

To identify the dominant contribution to the carrier emission, the photocurrent must be measured across a wide range of wavelengths and incident optical powers. To this end, we use the sample Design *I* in Figure 29 and Table 8, as this design allows to easily supply light directly on the MSM facet, without the need to correct the photocurrent amplitude for a wavelength-dependent silicon photonic waveguide. The coupling to such waveguide and the propagation in the waveguide would both introduce wavelength dependent losses. The amount of optical power coupled to the MSM test junction would strongly change while the wavelength is swept.
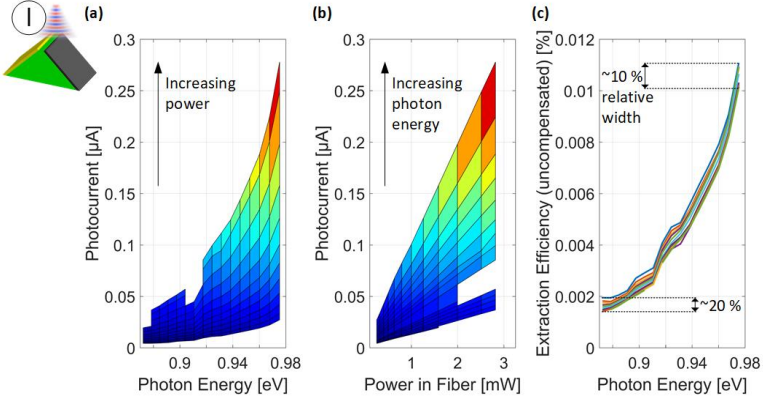
Figure 30. Photocurrent and extraction efficiency obtained from sample Design *I*, which relies on a wedge structure as indicated in the upper left corner. **(a)** The graph shows the photocurrent I_p (hf, P_{opt}) as an array of curves, with the photon energy hf as the variable. The optical power increases from the lowest curve to the highest. The power increments are not equidistant. The associated spacing is visible in the following subfigure. The photocurrent increases exponentially with the photon energy, which is a clear indication of Fowler photoemission. **(b)** The same data as (a), drawn with the optical power as the variable. The current essentially increases linearly with power. **(c)** The extraction efficiency features a band-like structure, which can be interpreted as a residual contribution of thermionic emission. The relative width constitutes an upper limit for this contribution and we estimate a maximum contribution between 10 % and 20 %.

We specify the measured photocurrents in relation to the photon energy (wavelength) given by the laser source and the optical input power in the fiber, as the actual power that reaches the MSM junction is unknown.

Figure 30 shows the photocurrent I_p (hf, P_{opt}) as an array of curves obtained from sample Design *I* as a function of photon energy hf and optical power P_{opt}. Figure 30 (a) shows that the photocurrent grows exponentially with the photon energy, which fits the expectation of Fowler's photoemission in Equation (23). The current grows by more than a factor of 6 in the measured photon energy range for all optical powers under consideration. Figure 30 (b) shows the same data set as a function of the optical input power. The current increases essentially linearly with power in the measured range, which is a further indicator of Fowler photoemission as the dominant carrier emission process.

The gaps in the measurement are due to performance limitations of our laser source, which did not stabilize in certain wavelength and power configurations.

To quantitatively estimate the current contribution of the thermionic emission from the heated carrier gas, we determine the extraction efficiency (EE), i.e., the number of electrons emitted per photon:

$$\mathrm{EE}\left(hf, P_{\mathrm{opt}}\right) = \frac{I_{\mathrm{p}}/\mathrm{e}}{P_{\mathrm{opt}}/hf} \tag{57}$$

In the case of pure Fowler photoemission, the EE should only depend on the photon energy and be otherwise entirely independent of the deposited power. All EE curves should come to rest on each other. However, a perfect overlay is not observed in Figure 30 (c). The EE curves still form a somewhat broadened band-like structure. This implies that there might be at least some contribution from thermionic emission. We evaluate the relative widths at the boundaries, i.e., the ratio of minimum current for a given photon energy and the maximum current. We find a 10 % relative width at the high energy threshold and a 20 % width at the lower one.

We assume that these numbers constitute only an upper limit for the contribution of thermionic emission to the overall current. Other unaccounted effects seem to play a role, too. This can be seen from the array of curves in Figure 30 (b), which are not perfectly linear and feature a slight downward slanting with the power, while thermionic emission would predict an upward slanting. One possible explanation could be that our model is incomplete, possibly omitting other nonlinearities in the dependency of the current and the optical power.

Evaluation of Barrier Height, Sample Design *II*

We proceed with measurements of the sample Design *II*. These samples have undergone practically the same fabrication process as the final PIPED designs. Specifically, all the etching steps that condition the metal-semiconductor interfaces are applied to this sample Design *II*, which was not the case for Design *I*. Due to this treatment and due to the possibility of coupling a wide range of wavelengths directly onto the MSM facet, this design is ideally suited
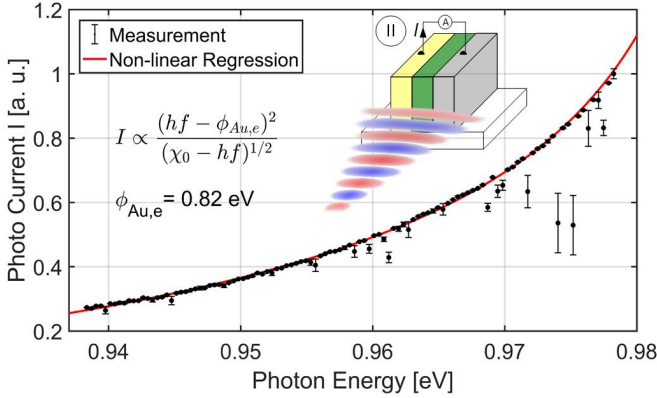
Figure 31. Measurement of Au-Si barrier height. A PIPED device of Design *II* is cleaved to have a length L ~ 10 mm, which is significantly longer than usual. The facet exposes the Au-Si-Ti junction to light with variable frequency *f*. The light is absorbed by the metals, and photoemission leads to a current *I* through the junction. At zero bias, only electrons from the Au contribute to the photocurrent. The measured photocurrent is normalized, and a nonlinear regression to Fowler's photoemission formula is applied, see Equation (23). The parameter χ_0 (total barrier height) is of no interest here. The barrier height $\phi_{Au,e} = 0.82$ eV (energy measured from the Fermi energy) can be extracted as a fitting parameter.

to investigate the barrier heights at the MS interfaces and to relate them to the final PIPED design.

The barrier heights at metal-semiconductor interfaces strongly depend on the materials and on the preparation of the surfaces, and values of around 0.3 eV [54] have been measured for interfaces of Au and neat Si surfaces. In our case, we rely on interfaces between Au and chemically treated Si surfaces prepared by reactive ion etching, which might significantly affect the actual barrier height, for example by gold diffusion or by surface states created by the Si processing [13]. To confirm the barrier height $\phi_{Au,e} = 0.82$ eV, which we used following [12, 13], we experimentally evaluated the photoemission current of sample Design *II*, which was exposed to a free-space irradiating wavelength-tunable laser beam, see Figure 31. We apply nonlinear regression of the measured current to fit the data to Fowler's photoemission model, see Equation (23). We extract the Au-Si barrier height as a fitting parameter. We find a barrier height at the Au-Si interface of $\phi_{Au,e} = 0.82$ eV, which is in good

agreement with [12]. Furthermore, we use the physical device simulator SILVACO ATLAS to model the dark current characteristics of the PIPED, see Appendix C.2. Using the parameters as described in Table 2, we find the predicted device currents to be in good agreement with our measurements. This is irrespective of whether the Si is slightly *p*- or *n*- doped, which is an indication that the barrier heights at the MS interfaces are indeed dominated by the conditioning during the reactive ion etching steps. These aspects lead us to believe that we employed a good approximation to the correct barrier height.

Wavelength Operating Range, Sample Design *III*

We assess the operating range of the PIPED by using sample Design *III* and measure the photocurrent for different wavelengths. Specifically, we chose 1550 nm (photon energy 0.8 eV) and 1300 nm (photon energy 0.955 eV), as both wavelengths are important for optical communications. Furthermore, from the choice of wavelengths, we expect a deeper insight to the emission characteristics in the PIPED: The photon energy of 1550 nm light is just below the largest barrier height in our device, i.e., the Au-Si interface $\phi_{\mathrm{Au,e}} = 0.82$ eV. The photon energy of 1300 nm light is significantly above this barrier height.

We apply an external bias voltage U_{DS}. The gate-gate source voltage U_{GS} is kept floating. The light is irradiated on a cleaved waveguide facet and then coupled to the PIPED. The same optical input power is used for all wavelengths. However, due to the coupling method, it is not known how much power ultimately reaches the PIPED after coupling to the waveguide and subsequently to the PIPED core. It is assumed that the difference in coupling efficiency for both wavelengths is not substantial. Figure 32 (a) shows the photocurrents for 1550 nm (I_{1550}) and 1300 nm (I_{1300}) obtained for a PIPED of 200 nm width and 5 µm length. The photocurrents I_{1550} and I_{1300} are in the same order of magnitude for both wavelengths. The same holds for the sensitivities S, consequently. We find that the general progression of the recorded traces is very similar, comprising a nonlinear increase of the current for large positive bias voltages and a saturation of the current for negative bias voltages.
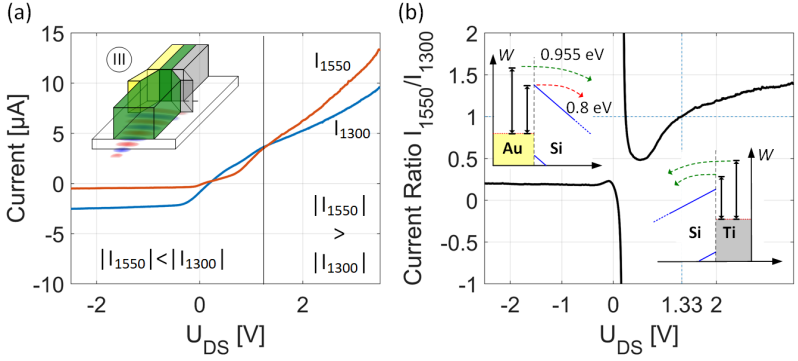
Figure 32. Performance comparison for two different operating wavelengths 1300 nm and 1550 nm. Sample Design *III* is used for this measurement. **(a)** The photocurrents I_{1550} and I_{1300} are in the same order of magnitude for both wavelengths. Two domains are discernible: For negative bias voltages U_{DS}, the current I_{1300} dominates, while I_{1550} dominates for large positive voltages. The reason for this switch can be understood from the current ratio, see Subfigure (b). **(b)** The ratio of the currents I_{1550} / I_{1300} changes very strongly with bias voltage. This is a consequence of the differences of photon energy and barrier heights. For negative bias voltages, the barrier at the Au-Si interface $\phi_{Au,e} = 0.82$ eV limits the carrier emission. The highest energetic carriers from absorption of 1300 nm light are well above this barrier (0.955 eV), while the highest energetic carriers from absorption of 1550 nm light barely come to rest at the level of the barrier (0.8 eV). For positive voltages, both photon energies are well above the barrier at the Si-Ti interface $\phi_{Ti,e} = 0.62$ eV. There is no substantial difference in carrier extraction efficiency for either wavelength, i.e., the current then only depends on the impinging photon rates. For the same optical input power, the beam of 1550 nm light contains 1.19 times the number of photons of the beam at 1300 nm. We find that the ratio of the currents tends towards $I_{1550} / I_{1300} = 1.4$, which is slightly larger than expected – possibly due to different absorption characteristics for the two wavelengths.

We find that the ratio of the sensitivities S_{1550} / S_{1300} changes very strongly with bias voltage, see Figure 32 (b): For negative bias voltages $|S_{1300}| > |S_{1550}|$ always holds. This condition changes for positive bias voltages. S_{1550} ultimately grows larger than S_{1300} for a positive bias of $U_{DS} > 1.33$ V. We interpret this finding as a consequence of the differences of photon energy and barrier heights. For negative bias voltages, the barrier at the Au-Si interface $\phi_{Au,e} = 0.82$ eV limits the carrier emission. The highest energetic carriers from absorption of 1300 nm light are well above this barrier, while the highest energetic carriers from absorption of 1550 nm light do not reach the level of the barrier. For positive voltages, both photon energies are well above the barrier at the Si-Ti interface $\phi_{Ti,e} = 0.62$ eV. Hence, there should be no substantial difference in carrier

extraction efficiency for either wavelength. The current then only depends on the impinging photon rates. For the same optical input power, the beam of light at 1550 nm features an $0.955 \text{eV}/0.8 \text{eV} = 1.19$ times larger photon rate than the beam at 1300 nm. We find that the ratio of the currents tends towards $I_{1550} / I_{1300} = 1.4$, which is slightly larger than the expected value. This could be due to differences in the absorption characteristics.

To generally confirm that the sensitivity increases with the wavelength under forward bias conditions (i.e., with electron emission from the Ti), we measured the wavelength range from 1300 nm until 1600 nm. We find that the sensitivity in fact increases in this range, confirming our conjecture. The measurable wavelength range was limited by the equipment that was available in our experiments.

In summary, we find that the useful operating wavelength range of PIPED at least extends from 1300 nm to 1600 nm. With the considerations of the section before, the high-energy cut-off threshold should be given by the Si absorption edge at 1130 nm (1.1 eV). This high-energy cut-off only applies for PIPED that are fed by silicon waveguides. These waveguides would consume all the optical input power before the light reaches the PIPED. Devices fed from free space would not feature any practical high-energy cut-off, as the silicon absorption is insignificant compared to the absorption in the metals. The low-energy cut-off threshold under forward bias should be given by the Si-Ti barrier height, predicting a still useful sensitivity at a wavelength of 2000 nm (0.62 eV). However, the ideal operating wavelength has ultimately not yet been determined and is up to future work.

Validation of Device Layout, Sample Design *III*

We have established in Chapter 4.2 that the key to the superior PIPED performance is the narrow silicon device core combined with the asymmetric design, using different metals on the drain and source electrodes. The asymmetric design leads to an efficient deposition of the optical power in the titanium electrode. The narrow core enables an efficient carrier extraction from the Ti, and tunneling through the metal-semiconductor barrier becomes possible when high external bias voltages are applied. These considerations have been the core principles in the PIPED design. This section shall supplement further experimental data, validating these design decisions.
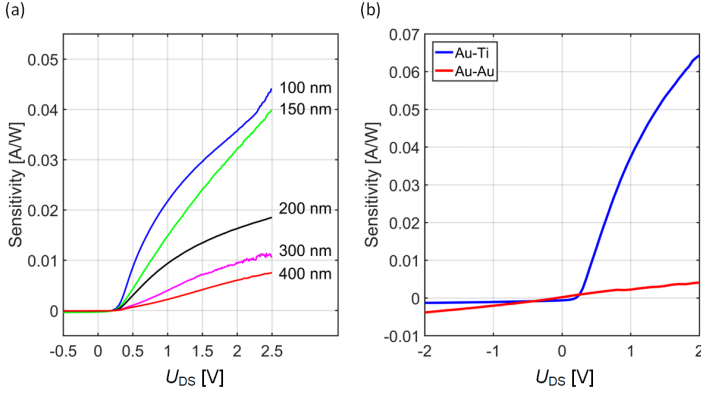
Figure 33. Performance comparison of PIPED with varying designs, measured at $\lambda = 1550$ nm. **(a)** Dependence of sensitivity on silicon core width. The array of curves shows the measured sensitivities of PIPED with core widths from $w = 400$ nm down to 100 nm. The sensitivity is the larger, the narrower the core width becomes. **(b)** Comparison of PIPED in the standard design (layer sequence Au-Si-Ti, blue curve) and an adapted symmetric design (Au-Si-Au, red curve). The symmetric Au-Si-Au design performs significantly worse at positive voltages compared to the standard design Au-Si-Ti, featuring only around 1/10 of its sensitivity.

Figure 33 (a) shows the sensitivities recorded from a set of PIPED, with core widths varying from $w = 400$ nm to 100 nm. The array of curves shows a clear trend, where narrower core widths lead to a better performance in terms of sensitivity. Specifically, while all PIPED show a moderate increase in sensitivity with bias voltage U_{DS}, the onset of tunneling is only visible for the narrowest PIPED with $w = 100$ nm in the recorded voltage range, see upwards slanting of the blue curve at $U_{DS} = 2.5$ V in Figure 33 (a). These results confirm that narrow PIPED are preferred to achieve a high sensitivity.

Figure 33 (b) compares the sensitivities of PIPED with the standard layer sequence Au-Si-Ti at the core to an adapted PIPED design with a symmetric sequence Au-Si-Au. To this end, the same chip design has been fabricated twice. One of the samples has been metallized with the standard layer sequence. For the other sample, the Ti deposition step has been replaced by another deposition of Au, thus resulting in a symmetric design. Unsurprisingly, the symmetric design yields a symmetric photocurrent characteristic around $U_{DS} = 0$. We also find that the adapted design performs significantly worse

compared to the standard design in the high positive bias voltage range, which is the relevant one for the detector operation. The adapted design features less than 1/10 of the performance of the standard design. Interestingly, in the otherwise unimportant reverse-bias voltage range, the symmetric Au-Si-Au design performs better. We attribute this to the capability of Au to both emit electrons on the negatively charged side and to emit holes on the positively charged side. Note that the samples in Figure 33 (b) are not identical in core widths. The symmetric design has a width of $w = 150$ nm, the asymmetric design has a width of $w = 125$ nm. The set of usable samples was limited on this fabrication run due to a generally low yield. However, we are convinced that the choice of samples is not detrimental to our findings in a substantial way.

In summary, we confirm that the PIPED design with an asymmetric core layer sequence Au-Si-Ti and a narrow core width provides the optimal performance.

C. PIPED Numerical Modelling and Device Design Parameters

Numerical device modelling is an essential prerequisite to understand the device physics on a fundamental level. For the purposes of photodetection, the photonic and electronic properties of PIPED are of special importance. We numerically simulate the PIPED to identify the key design parameters. Based on this, we used the improved understanding to define design guidelines and an optimized device layout in the Chapters 4 and 5. The following sections explain the details of the numerical modelling environments, as well as their capabilities and limitations.

C.1 Nanophotonic Modelling with COMSOL

We use COMSOL Multi Physics as a simulation environment for the nanophotonic modelling of PIPED. This tool offers a broad set of simulation methods, from which we employ a fully vectorial 3D finite-element simulation in frequency domain. The goal of the simulation is to analyze the coupling of light to the PIPED and to optimize the absorption characteristics within.

The simulated device has been structured to represent the actual fabricated PIPED device. The simulation domain contains the silicon-photonic feed waveguides, the tapered photonic-to-plasmonic mode converter, and the detector core itself. The structural features have been taken from the device design and verified by scanning electron microscopy. Figure 34 shows a comparison of the actual device, (a), and a graphical representation of the simulated structure, (b). The length of the plasmonic taper L_{Taper} and the detector core width w are used as simulation parameters that have the most crucial influence on the PIPED optical performance. All other structural parameters are kept constant. Table 10 shows a list of the parameter values typically used in our simulations. We define the permittivities of the materials that make up the PIPED (i.e., silicon, gold [17], titanium [18], and silicon oxide) according to published reference values for the metals, which are strongly frequency-

dependent. For the dielectrics, we use the predefined permittivity values of COMSOL. Table 9 lists the permittivities employed in our simulations for the relevant wavelength $\lambda = 1550$ nm.
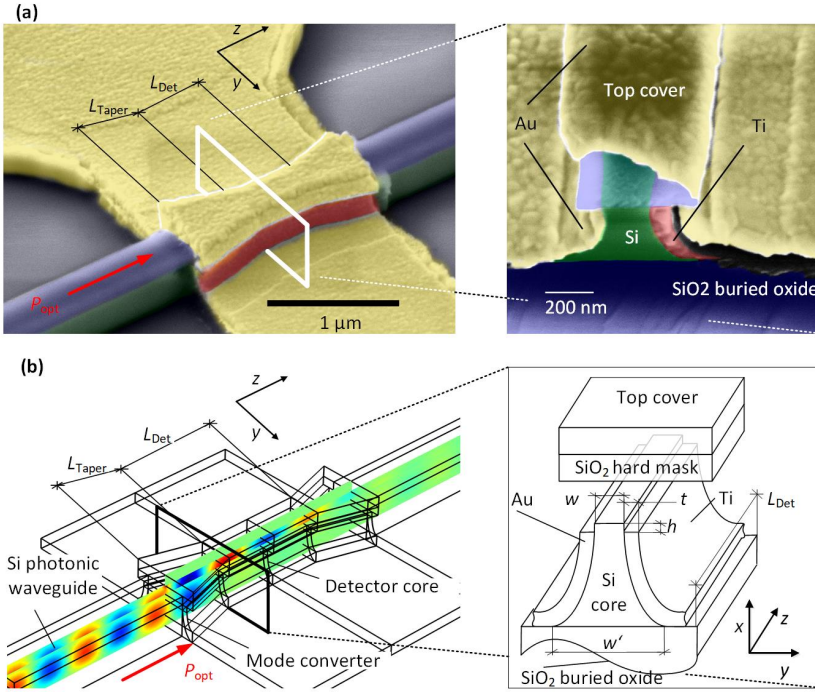
Figure 34. Structural comparison of a fabricated PIPED and the simulated device in COMSOl Multiphysics. **(a)** PIPED micrographs of top view (left) and cross-section (right). L_{Taper} and L_{Det} denote the lengths of the photonic-to-plasmonic mode converter and the detector core. **(b)** Schematic of the simulated PIPED structure with viewing angles similar to (a). The relevant free design parameters are highlighted. The typical parameter values are given in Table 10. The structure has been defined to approximate the fabricated PIPED as closely as possible

Table 9. Optical properties of the PIPED photonic model.

Material	Relative permittivity at $\lambda = 1550$ nm	Source
Gold	$\varepsilon_{Au} = -112.7 - j\,6.85$	[17]
Titanium	$\varepsilon_{Ti} = 1.6 - j\,30.56$	[18]
Silicon	11.7	Comsol predefined material property
Silicon Oxide	2.09	Comsol predefined material property

Table 10. Structural parameters of the PIPED photonic model. Parameters as defined in Figure 34.

Parameter	Value range	Comment
L_{Taper}	150 – 800 nm,	Length of mode converter
L_{Det}	0.5 – 20 µm	Length of detector core
w	50 - 400 nm	Narrowest Si diameter at the apex
w'	190 – 540 nm	Widest Si diameter at the base. Typically $w + 2\times120$ nm. 120 nm is the lateral etch depth.
h	20 – 50 nm	Bare silicon height
t	50 nm	Metal thicknesses
Si phot. wg. width	550 nm	/
Si phot. wg. height	300 nm	Value fixed by SOI layer sequence
SiO$_2$ hard mask width	/	Constrained by $w + 2\times120$ nm. 120 nm is the lateral etch depth.
SiO$_2$ hard mask height	80 nm	
SiO$_2$ buried oxide thickness	2000 µm	Value fixed by SOI layer sequence

The light enters the simulation domain in the silicon photonic input waveguides via ports. Boundary mode analyses are applied on the ports that provide the photonic waveguide modes and the associated effective propagation indices n_{eff}. The input waveguides are capable of guiding modes with a dominant horizontal electric field component (henceforth TE mode) and a dominant vertical electric field component (henceforth TM mode). We employ the boundary mode analysis for the input and the exit waveguide such that the scattering parameters S_{12} (power transfer through PIPED) and S_{11} (reflected power from PIPED) can be extracted. We use the same TE/TM terminology to refer to PIPED modes. This is because the TM mode in the silicon photonic waveguide couples mainly to the PIPED mode with dominant vertical field component, and analogously, the TE mode couples mainly to the PIPED mode with dominant horizontal field component.

The simulation yields the electric field distribution $E(x,y,z)$ across the entire simulation domain. To extract figures for the optical coupling efficiency and the absorption characteristics, we evaluate the local light absorption, which is in proportion to $\text{Im}(\varepsilon) \times |E|^2$, where only the metals contribute to the absorption. We analyze the taper and the detector core separately. We also take the metallic subsections (i.e., the parts made of Au, or the parts made of Ti) into account separately. Integrating the local absorption over the volume of metallic subsections yields the total absorption in this subsection. We end up with the optical power P_{opt} deposited in the Au part of the taper, the Ti part of the taper, the Au part of the core and the Ti part of the core.

Note that analyzing the evolution of the Poynting $\vec{S}$ vector in the device is an alternative method for the absorption evaluation. The modulus of $\vec{S}$ will decrease by absorption in the direction of propagation. Integrating the z-component of the Poynting vector across the (x,y) plane will result in the optical power at a given depth z in the device $P_{opt}(z) = \iint \vec{S}\big|_{z=\text{const}} \,\mathrm{d}x\mathrm{d}y$. As the absorption $\text{Im}(\varepsilon) \times |E|^2$ and the residual power $P(z)$ in the device are connected to each other, evaluation of the $P_{opt}(z)$ is a helpful tool for a sanity check.

To analyze the eigenmodes of the PIPED, the simulation domain can be restricted from 3D to 2D, only including a cross-section of the detector core in the x-y plane. In this reduced domain, COMSOL can simulate the eigenmodes and the associated complex-valued effective mode indices $\underline{n}_{eff} = n_r - \mathrm{j}n_i$, the

imaginary part of which defines the absorption coefficient. We find that more than 99% of the incident light is absorbed in a PIPED with a length of only $L_{Det} = 5$ µm and for all technically relevant taper lengths L_{Taper} and core widths w, see Appendix A.2 .
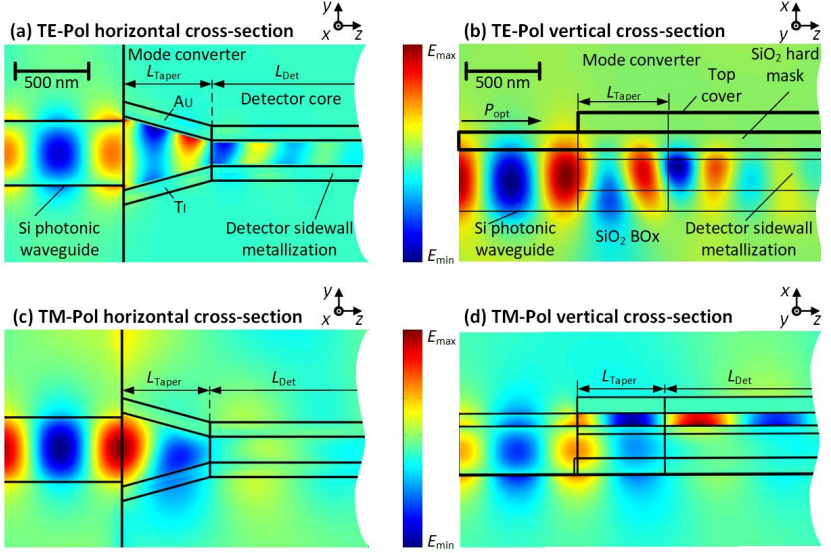
Figure 35. PIPED internal field distribution for the TE and TM modes. The associated cut planes of the cross-sections intersect at the center of the PIPED. **(a)** TE-polarization and horizontal cross-section (top view). The light enters the simulation domain in silicon photonic input waveguides. The photonic mode couples to the metallic sidewalls of the PIPED in the tapered mode converter. The light is then coupled to the detector core. The length of the mode converter is L_{Taper} and has a strong influence on the coupling efficiency. The light absorption in the mode converter and in the detector is high, and the field strength quickly decays. **(b)** Same mode profile as in (a) as a vertical cross-section (side view). The light concentrates at the apex of the silicon core, where the diameter is the narrowest. The subfigures **(c)** and **(d)** show the same cross-sections for the TM polarization. The TM mode does not couple to the taper. Instead, the light couples to an SPP propagating along the metallic top cover and concentrates in the SiO2 hard mask.

Most of the results for the photonic modelling have already been covered in the Chapters 4 and in Appendix A.2 and are not repeated here. We supplement this data by the longitudinal mode profiles of the PIPED for the TE and the TM polarization, shown in Figure 35. The actual taper length L_{Taper} and the actual detector core width w have no strong influence on the general appearance of the longitudinal mode profiles, and the results discussed below are typical for the entire respective parameter ranges.

Figure 35 (a) and (b) show the field distribution for TE excitation in the y-z-plane and the $(x$-$z)$ plane, respectively. Both planes intersect at the center of the PIPED, i.e., at half its height and half its width. It can be seen in (a) that the light concentrates between the Au and the Ti electrode in the detector core and in (b) that practically no light resides in the SiO_2 hard mask layer. The light is quickly absorbed while penetrating deeper into the device along the z-direction.

The TM polarization behaves differently: Essentially no light enters the detector core. Instead, the light couples to the Au top cover and enters the SiO_2 hard mask layer. The corresponding electric field distributions for the TE and the TM modes $E(x, y)$ can be found in Chapter 4.5, Figure **11**. For both input polarizations, the lateral mode profiles are very similar to the eigenmodes of the PIPED.

C.2   Electronic Modelling with SILVCACO ATLAS

We use the 2D-simulation environment SILVACO ATLAS (https://silvaco.com/) to model the electronic properties of PIPED and FE-PIPED, respectively. ATLAS is a finite-element physical device simulator that is built upon the drift-diffusion model and solves the equations of continuity (Eq. (17) and Eq. (18)) and Poisson's equation for any given semiconductor device structure. This allows to simulate the potential $\varphi(x,y)$, the carrier densities, and the current densities throughout the device at arbitrary biasing conditions. We use ATLAS to simulate the above device characteristics in an $(x$-$y)$ cross-section of the detector core, i.e., in a cut plane whose normal vector is perpendicular to the photocurrent and parallel to z, the direction of light propagation. The 2D-simulation has the advantage to require much less computational power than a full structural model in three dimensions would.

The tapered mode converters are neglected, as they cannot be accommodated in the 2D simulation. We verified, however, that their current contribution, at least for the dark current, is negligible. Currents have the units of A/µm due to the restriction of the simulation to a cut-plane. ATLAS assumes a device length of 1 µm to output the current in Amperes. Comparability to measured currents from actual devices is given by relating the measured currents to the device length.

ATLAS employs the *Universal Schottky Tunneling* model (*UST*) to simulate the current across a MS interface. The model includes the effects of voltage-dependent tunneling and thermionic emission. However, the *UST* model is only suitable to simulate the dark-current characteristics of the FE-PIPED and is not able to make quantitative predictions for the sensitivity S_0 for devices based on internal photoemission. The limitation arises from a boundary condition of the model, which always assumes an electron energetic distribution $f(W)$ at a given temperature in thermal equilibrium with its surroundings. This model is hence unable to take into account out-of-equilibrium distributions with hot electrons or holes.

There is currently no verified algorithm commercially available that would be capable to fully simulate S_0, neither by employing empiric models nor by considering first principles. This is due to the limited knowledge about the energetic distribution $f(W)$ of the hot carriers in the metals, which crucially defines the carrier transmission probability across the barrier.

The simulated device cross-section mirrors the actual FE-PIPED devices as fabricated. We define the Au electrode as the drain contact, the Ti electrode as the source contact and the Au top cover as the gate contact. The associated voltages between the electrodes are U_{DS} (drain-source) and U_{GS} (gate-source). The influence of the gate voltage has been discussed in the previous chapters. The results discussed in this chapter are generally valid with and without an applied gate voltage. For the sake of simplicity, we set $U_{GS} = 0$ in the following.

The detector core width w is chosen according to the range used in our experiments. The definition of the mesh is crucial for obtaining reliable results for the currents. To this end, we use a mesh density of 0.5 nm by 0.5 nm in the silicon close to the apexes of the metals, where tunneling is most prominent,

and just below the gate oxide, where the conductive carrier channel forms. We use a wider mesh in other regions with little or no contribution to the current to limit the computational load. We have verified that these simplifications do not impair the numerical results.

The barrier heights at the Au-Si and the Si-Ti interfaces are indirectly controlled for by the differences of the metal work functions ϕ_W and the Si electron affinity χ_{Si}. For the Au-Si interface this exemplarily reads $\phi_{Au,e} = \phi_{W,Au} - \chi_{Si}$. We chose to use the predefined value for χ_{Si} and use $\phi_{W,Au}$ and $\phi_{W,Ti}$ as free parameters. During our exploration of the parameter ranges, we found that the difference of the work functions (and consequently the difference of the barrier heights) $\phi_{W,Au} - \phi_{W,Ti}$ defines the flat-band bias voltage φ_{bi} and hence the threshold voltage which must be exceeded to create any significant current flow between
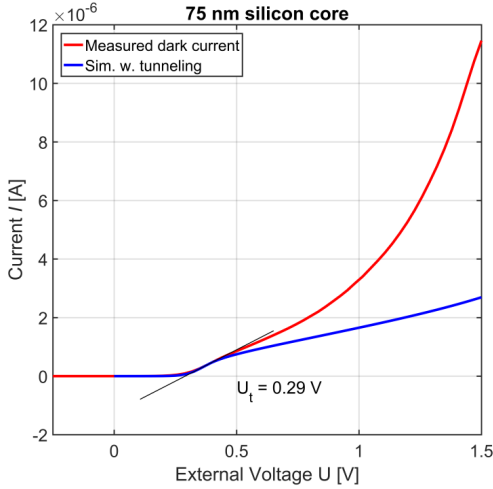
Figure 36. Comparison of measured (red) and simulated (blue) dark current-voltage characteristic for a PIPED with a core width of $w = 75$ nm. We approximate the measured device current I by a linear function around the flat band voltage U_t. The U-intercept $\varphi_{bi} = U_t$ is identical to the flat band voltage and to the difference of the work functions $\varphi_{bi} = (\phi_{W,Au} - \phi_{W,Ti})/e$. The simulation is able to reproduce the actual device current very well for the voltage range up to 0.5 V. For larger voltages, an exponential increase of the current with the applied voltage can be observed, which marks the onset of tunneling. The curves grow at different rates. The simulation is not able to fully cover the tunneling domain.

the drain and the source contact. The absolute values of ϕ_W set the slope of the current growth for voltages $U > \varphi_{bi}$ and before tunneling sets in.

We use the in-built optimization algorithm of ATLAS to fit the simulated currents to the measured currents of PIPED. To this end, we approximate the device current I by a linear function $I = m \times U + b$ around the flat band voltage U_t. The U-intercept $\varphi_{bi} = U_t$ is identical to the flat band voltage and to the difference of the work functions $\varphi_{bi} = -b/m = (\phi_{W,Au} - \phi_{W,Ti})/e$. We optimize $\phi_{W,Au}$ until the current slopes m are in good agreement with each other. Figure 36 shows the result of such optimization for a PIPED with a detector core width of $w = 75$ nm simulated at room temperature. The measured dark current is plotted in red and the associated simulation is plotted in blue. The two curves agree generally well with each other for small voltages $U < 0.5$ V. The resulting barrier height difference for this particular device is $\varphi_{bi} = 0.29$ V. We find that φ_{bi} varies for different samples from different fabrication campaigns, but is generally the same for a distinct set of samples from the same fabrication run. For example, the PIPED discussed in Chapter 4 features a slightly smaller $\varphi_{bi} = 0.2$ V than the one shown in Figure 36. We attribute this variation to differences of the barrier heights that form as a consequence of the silicon processing. With the prior analysis, we understand φ_{bi} as a viable predictor of the barrier height difference for a particular PIPED.

We further find that the parameter sets that predict the current fairly well around the flat-band voltage do not fit anymore for larger voltages $U > 0.5$ V. Here, the tunneling current dominates and this current is obviously underestimated by our simulation, see Figure 36. We can artificially increase the exponential growth of the current by running the simulation at a higher device temperature, providing a better match for the tunneling currents. However, in this case, the slope m around the flat-band bias voltage does not fit anymore well enough. The reason for this finding might be a rapid heating of measured PIPED, leading to higher dark currents as the devices are under test. In addition, the Richardson constants for one or both interfaces might not match well with the actual conditions at the interfaces, leading to an underestimation of the tunneling current at high voltages.
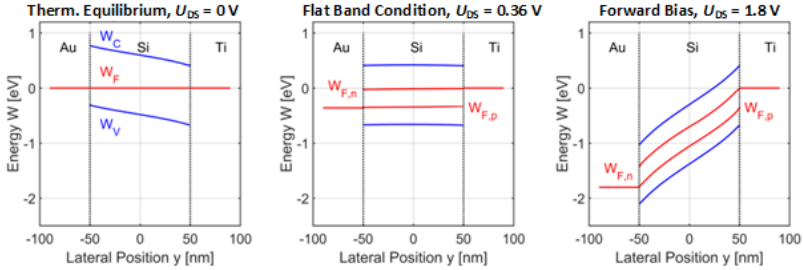
Figure 37. Drain-source voltage dependence of PIPED band bending ($U_{GS} = 0$ V). The figure displays a sequence of biasing conditions, with **(a)** $U_{DS} = 0$ V, **(b)** 0.36 V and **(c)** 1.8 V. The band edges of the conduction band W_C and the valence band W_V are mostly parallel throughout all biasing conditions. The quasi Fermi levels $W_{F,n}$ and $W_{F,p}$ split when a bias is applied and reach a maximum separation of 0.36 V. This is identical to the flat band voltage for this particular device setup.

We further investigate the band bending and its voltage-dependence for a PIPED with a hypothetical barrier height difference of $(\phi_{Au,e} - \phi_{Ti,e}) = 0.36$ eV and with a Au-Si barrier height of $\phi_{Au,e} = 0.82$ eV. The conduction band edges W_C and the valence band edges W_V of the Au-Si-Ti junction are extracted from the simulation results at the apex of the detector core, i.e., where the silicon is the narrowest. The results are displayed in Figure 37.

The figure shows a sequence of biasing conditions, with (a) $U_{DS} = 0$ V, (b) 0.36 V and (c) 1.8 V. The band edges are mostly parallel throughout all biasing conditions. The quasi Fermi levels $W_{F,n}$ and $W_{F,p}$ split when a bias is applied and reach a maximum separation of 0.36 V, which is identical to the flat band voltage. The band diagrams are in good agreement with the assumed model in the previous chapters. The flat-band voltage hence is a viable predictor of the band alignment in the silicon.

D.   Dynamic Device Characteristics

The FE-PIPED has the potential to reach unprecedented optoelectronic bandwidths, outperforming conventional Si/Ge photodetectors by an order of magnitude. We use time-domain numerical simulations of the intrinsic device dynamics of the PIPED to assess the device's ultimate performance limits. To this end, a device model is created that reflects the device's physical properties. A time-dependent broad-band photocurrent is created in the device as a stimulus to probe the device dynamic behavior. The external current response to this stimulus is measured as a function of the gate-source voltage U_{GS}. The opto-electronic bandwidth is inferred by relating the dynamics of the external current and the stimulus. The following sections explain the methodological approach and discuss the results.

D.1   Methodological approach

The dynamic device response of the FE-PIPED manifests itself in the frequency-dependent amplitude of its photocurrent. The temporal response in the external current is determined by the timing of the hot carrier creation and transport within the device, see Sections 4.2 and 5.2 for details. The largest fraction in time is required by the carrier transfer from the injecting contact to the opposing contact at the far end of the device. The high-speed current amplitude will generally decrease with increasing driving frequency f, as the device is less capable to follow the time-harmonic stimulus. The gate voltage U_{GS} is key to maximizing the opto-electronic bandwidth by controlling the carrier density in the FE-PIPED, where the device dynamics is determined by dielectric relaxation instead of the comparably slower carrier transit time through the device. However, the gate voltage also increases the injection efficiency of carriers from the metal into the silicon. To address the different effects of the control voltages independently, the opto-electronic sensitivity S is separated in two factors:

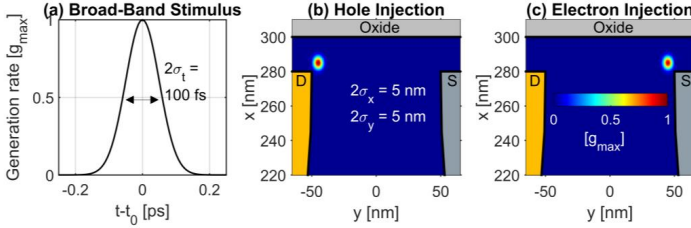$$S(f, U_{\text{DS}}, U_{\text{GS}}) = S_0(U_{\text{DS}}, U_{\text{GS}}) H(f, U_{\text{GS}}) \tag{58}$$

Figure 38. High-speed stimulus. The generation rate g is defined as a Gaussian function in time and space to place test charges in the FE-PIPED and to observe the resulting response of the device current. **(a)** Timing of broad-band stimulus. The generation rate assumes the form of a narrow Gaussian. The width is defined by the standard deviation $2\sigma_t = 100$ fs. This very short impulse provides a broad spectrum in the frequency domain with spectral components well above 10 THz. **(b)** Hole injection close to the Au drain contact. Electron-hole pairs are created in a narrow region adjacent to the apex of the Au. The center-coordinates of the Gaussian distribution are $x_0 = 285$ nm and $y_0 = -45$ nm. The widths of the Gaussian are $2\sigma_{x,y} = 5$ nm. The electron-hole pair separates under forward bias conditions $U_{\text{DS}} > 0$. The electron is immediately absorbed by the Au contact and a hole remains. The hole is subject to the electric field in the device and traverses towards the Ti source contact, which is grounded. Relating the frequency components of the current response measured at the source and drain terminals to the stimulus allows inferring the transfer function of the FE-PIPED in case of a dominant hole photocurrent. **(c)** Electron injection close to the Ti drain contact, $y_0 = 45$ nm. Analogously to (b), the electron-hole pair separates, and the hole immediately vanishes into the Ti. The electron remains. This configuration is used to assess the transfer function of the FE-PIPED in case of a dominant electron photocurrent.

$S_0(U_{\text{DS}}, U_{\text{GS}})$ is the DC sensitivity, which models the voltage dependence of the carrier injection. It can be assumed that the hot-electron dynamics and the injection in the metals is approximately independent of the driving frequency and the same applies to S_0. The frequency dependence of the current transfer through the device is entirely modelled by $H(f, U_{\text{GS}})$, which is the opto-electronic transfer function.

We analyze the transfer function H with the help of numerical simulation. We use the 2D-simulation environment provided by SILVACO ATLAS to create the FE-PIPED device model and to execute the simulation. ATLAS is a physical device simulator that is built upon the drift-diffusion model and solves the equations of continuity (Eq. (17) and Eq. (18)) for any given semiconductor device structure. The details of the modelling are explained in Appendix C.2.

The available numerical tools are currently limited in their capability to model internal photoemission, which also applies to ATLAS. We circumvent this limitation by using a specific method to model an optical stimulus: We approximate an injected hot charge carrier by placing test charges in close proximity to the Au or the Ti contact, see Figure 38 (b) and (c). We choose positions close to the metal apexes, where the electric field is the strongest and tunneling would create the highest carrier injection rate.

To inject the test charges, we define a time- and space-dependent generation rate $g(t,x,y)$ of electron-hole pairs in the silicon next to the apex of the Au source or the Ti drain contact. Under forward biasing conditions $U_{DS} > 0$ V, an electron created close to the Au will immediately be absorbed by this contact and a hole remains. The remaining hole is subject to the same electric field as an injected hot hole would be, and a current is created between the device terminals. Hence, it is fair to assume that the remaining hole approximates the dynamic behavior of an optically generated hot hole that overcame the barrier at the MS interface and that was injected to the silicon.

An electron-hole pair created at the Ti contact behaves analogously: The hole vanishes into the metal and the electron contributes to the current through the device. The definition of an electron-hole pair instead of only a single charge of interest is a constraint imposed by the ATLAS simulation environment. However, it is assumed that the additional charge does not significantly affect the simulation results, as it immediately vanishes into the metal contact.

The generation rate g is defined as a Gaussian function both in time and space. The temporal and spatial widths are $2\sigma_t = 100$ fs, $2\sigma_{x,y} = 5$ nm. Figure 38(a) illustrates the Gaussian impulse. Due to its short temporal width, the impulse features frequency components well above 10 THz and is specifically chosen to provide a broad-band stimulus of the FE-PIPED. The center of the electron-hole pair generation is chosen at $x_0 = 285$ nm and $y_0 = -45$ nm close to the Au contact or at $y_0 = 45$ nm close to the Ti contact, see Figure 38 (b) and (c). This definition matches the location where ATLAS' *Universal Tunnelling Model* provides the largest generation rate due to tunneling, generated by a strong field enhancement at the sharp metal apexes. It is fair to assume that tunneling currents should behave similarly for dark and photocurrents alike, which justifies our definition of the electron-hole pair

generation. The widths $2\sigma_{x,y}$ are designed to be large enough to avoid numerical instabilities, while at the same time still being sufficiently small compared to the silicon core width w that needs to be traversed. This avoids detrimental influences on the simulation results. The peak magnitude of the generation rate is defined as $g_{max} = 10^{25}$ cm^{-3}s^{-1}. This constitutes an empirical value which results in a measured photocurrent in the order of 1 µA, which matches the typical photocurrents of the real PIPED.

We analyze the dynamic behavior of the created electrons and the holes separately. This means that the electron-hole pair is either placed at the Au drain contact for simulation of a hole transfer, or at the Ti source contact for simulation of an electron transfer. We do not place test charges at both metals at the same time. This is justified, as our results indicate that for most operating conditions, only one carrier type dominates the photocurrent: Holes dominate for $U_{GS} < 0$ V while electrons dominate for large positive gate voltages $U_{GS} > 4$ V, see Section 5.5.1.

D.2   Impulse Response

We predict the impulse response of the PIPED by exciting the biased device model with the broad-band stimulus shown in Figure 38 (a). We record the time-dependent excess currents due to the stimulation at the device terminals I_{src} (source), I_{drn} (drain) and I_{gate}. By this definition, the dark currents are subtracted, and the remaining current can be fully attributed to the optical stimulation. The drain-source voltage is fixed at $U_{DS} = 1.5$ V. This choice ensures reaching the saturation velocity of drifting carriers for the given device of width $w = 100$ nm. The gate voltage is swept in a range of $U_{GS} = -10$ V to $U_{GS} = 10$ V.

We execute the simulation of the impulse response separately for holes and for electrons by choosing the location of the carrier generation according to Figure 38 (b) or Figure 38 (c). These modes of excitation are in analogy to a strongly absorbing *p-i-n* photodiode, which is stimulated by a short light pulse close to the drain contact, (b), or close to the source contact, (c). However, while the drift region of the strongly absorbing *p-i-n* photodiode is depleted of mobile charges, the PIPED charge density depends on the gate voltage U_{GS}. In the following, we investigate how the carrier densities impact the PIPED device

dynamics. Figure 39 shows examples of the impulse response of I_{src}, obtained for hole and electron injection and for gate voltages of $U_{GS} = 0$, -10 V and 10 V, respectively. The width of the FE-PIPED is $w = 100$ nm.

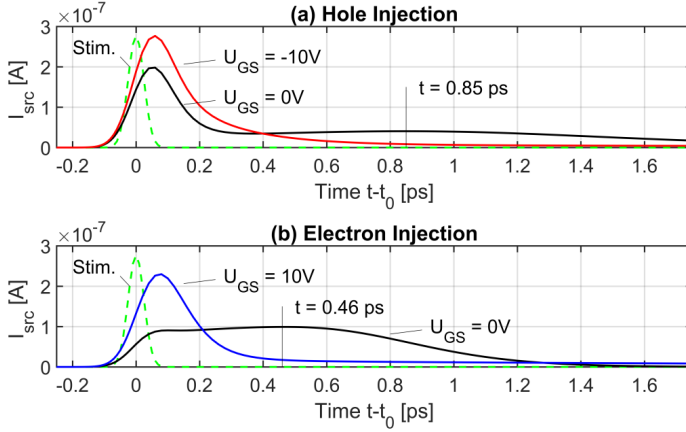**Figure 39.** Impulse response of $I_{src}(t)$ for a FE-PIPED with a width of $w = 100$ nm. The drain-source voltage is set to $U_{DS} = 1.5$ V to ensure that carriers drift at their saturation velocity. The dashed green curve illustrates the stimulus as a reference, which is drawn in arbitrary units. The time axis has been gauged according to the peak of the stimulus which is at $t - t_0 = 0$. **(a)** Hole Injection. The black curve depicts the dynamics of the device at zero bias on the gate. The current comprises two distinct domains. A first current peak arises at $t - t_0 = 60$ fs shortly after the stimulus. This peak appears more than ten times faster than expected when hole drift at the saturation velocity is assumed. The full width is $\sigma = 140$ fs, measured at 61 % of the maximum, defined in analogy to the width of the stimulus. We attribute this ultra-fast response to dielectric relaxation sustained by the hole channel under the gate oxide, which is already present at zero gate voltage. This first peak is followed by a second tail with a very broad and much slower current spanning more than 1 ps. This tail has its maximum at $t - t_0 = 850$ fs. We attribute this broader current to holes drifting through the Si instead of taking part in dielectric relaxation. Assuming a saturation drift velocity of 10^7 cm/s, a time of 850 fs corresponds to a drift distance of 86 nm, which agrees acceptably well with the device width $w = 100$ nm. We understand the broad shape of the current as a consequence of the statistical nature of the carrier drift. The red curve shows the current at a high negative gate voltage of $U_{GS} = -10$ V, which strengthens the p-channel. The long tail seen at zero gate voltage vanishes and the first peak grows in strength. In this operating condition, most of the injected carriers contribute to the dielectric relaxation current and the speed of the device is significantly increased by the gate voltage. **(b)** Electron injection. The current response has a broad appearance at zero gate voltage. This is similar to the behavior of holes. We understand this as a large fraction of electrons drifting through the device. The current reacts significantly faster at $U_{GS} = 10$ V and no relevant difference to the hole current at the same gate voltage can be recognized.

The black curve in Figure 39 (a) depicts the time-dependent current of the device for hole injection and at zero bias on the gate, $U_{GS} = 0$ V. The current comprises two distinct domains. A first current peak arises at $t - t_0 = 60$ fs shortly after the occurrence of the stimulus. A second peak follows with a reduced amplitude and a much larger width spanning more than 1 ps. We associate each peak with a different carrier transport mode in the Si.

We attribute the first peak to dielectric relaxation due to the interaction of the injected holes with the hole channel under the gate oxide. This channel is already present at $U_{GS} = 0$ V, albeit weakly pronounced. The peak occurs more than ten times faster than expected from holes drifting at the saturation velocity. The full width is $\sigma = 140$ fs. The broadening of the first peak compared to the stimulus can be explained by the variation of the carrier density in the channel. This density decreases from the gate oxide into the bulk Si, hence affecting the local dielectric relaxation time. Each part of the channel introduces a slightly different time delay.

We attribute the broad second peak to holes drifting through the bulk Si outside of the conductive channel. This peak has its weakly pronounced maximum at $t - t_0 = 850$ fs. Assuming a saturation drift velocity for holes in Si of $v_{s,h} = 1 \times 10^7$ cm/s, a transit time of 850 fs corresponds to a drift distance of 86 nm. This value agrees well with the device width w, considering that the injection location is shifted slightly into the Si. We attribute the broad appearance of the tail to a strong dispersion in the device, i.e., the drift velocities in the device differ strongly between the individual carriers.

Due to the design of our injection model, we would have expected a further current peak due to the minority carriers that immediately vanish into the drain electrode. However, this additional peak is not visible. We assume that this peak occurs too fast for the simulation to follow.

We now apply a gate voltage of $U_{GS} = -10$ V, strengthening the p-channel. The red curve in Figure 39 (a) shows that the carrier transport mode has significantly changed. The slow tail that existed without a gate voltage effectively vanished. We understand this as a consequence of the larger carrier density in the channel, which also reaches farther into the device. In this operating condition, most of the injected carriers interact with the hole channel, and the carrier transit time

does not play a role anymore. This change of carrier dynamics significantly increases the speed of the device.

Figure 39 (b) depicts the dynamics of the device for electron injection. The characteristic for zero gate voltage, see black curve marked with $U_{GS} = 0$ V, shows a distinctively different behavior as compared to the hole injection at zero voltage in Figure 39 (a). It appears as though the two peaks merged into one. Two distinct "bumps" can still be identified. Our interpretation of this shape is similar as before: The first bump corresponds to the interaction of the injected carrier with the channel under the gate oxide. However, the strength of the current and the speed are reduced. This is because now minority carriers interact with the majority carrier channel. The second bump at $t - t_0 = 460$ fs corresponds to electrons drifting through the Si. The blue curve shows the impulse response at $U_{GS} = 10$ V. Applying a strong positive gate voltage inverts the semiconductor and creates an n-channel. Injecting electrons leads to a very fast dielectric relaxation in the channel. The timing characteristics are now similar to the corresponding situation in Figure 39 (a), and the carrier transit outside of the channel is effectively eliminated.

In conclusion, we have shown that applying a gate voltage allows to control the carrier dynamics in the FE-PIPED. In particular, both positive and negative gate voltages lead to a significantly shorter impulse response as compared to the case without an applied gate voltage. We attribute the increase in device speed to the increased dielectric relaxation in the channel when gate voltages are applied.

E.   Bibliography

[1]   D. A. B. Miller, "Device Requirements for Optical Interconnects to Silicon Chips," Proceedings of the IEEE **97**, 1166-1185 (2009).

[2]   J. Leuthold, C. Hoessbacher, S. Muehlbrandt, A. Melikyan, M. Kohl, C. Koos, W. Freude, V. Dolores-Calzadilla, M. Smit, I. Suárez, J. Martínez-Pastor, E. P. Fitrakis, and I. Tomkos, "Plasmonic Communications: Light on a Wire," Optics and Photonics News **24**, 28-35 (2013).

[3]   M. Bohr, "14 nm Process Technology: Opening New Horizons (http://www.intel.com/content/dam/www/public/us/en/documents/pdf/foundry/mark-bohr-2014-idf-presentation.pdf)" (Intel, 2014), retrieved 15 March 2016.

[4]   S. Muehlbrandt, A. Melikyan, T. Harter, K. Köhnle, A. Muslija, P. Vincze, S. Wolf, P. Jakobs, Y. Fedoryshyn, W. Freude, J. Leuthold, C. Koos, and M. Kohl, "Silicon-plasmonic internal-photoemission detector for 40 Gbit/s data reception," Optica **3**, 741-747 (2016).

[5]   S. Muehlbrandt, T. Harter, C. Füllner, S. Ummethala, S. Wolf, A. Bacher, L. Hahn, M. Kohl, W. Freude, and C. Koos, "Field-effect silicon-plasmonic photodetector for coherent T-wave reception," Optics Express **29**, 21586-21602 (2021).

[6]   T. Harter, S. Muehlbrandt, S. Ummethala, A. Schmid, S. Nellen, L. Hahn, W. Freude, and C. Koos, "Silicon–plasmonic integrated circuits for terahertz signal generation and coherent detection," Nature Photonics **12**, 625-633 (2018).

[7]   T. Harter, S. Ummethala, M. Blaicher, S. Muehlbrandt, S. Wolf, M. Weber, M. M. H. Adib, J. N. Kemal, M. Merboldt, F. Boes, S. Nellen, A. Tessmann, M. Walther, B. Globisch, T. Zwick, W. Freude, S. Randel, and C. Koos, "Wireless THz link with optoelectronic transmitter and receiver," Optica **6**, 1063-1070 (2019).

[8]   T. Harter, C. Füllner, J. N. Kemal, S. Ummethala, J. L. Steinmann, M. Brosi, J. L. Hesler, E. Bründermann, A. S. Müller, W. Freude, S. Randel, and C. Koos, "Generalized Kramers–Kronig receiver for coherent terahertz communications," Nature Photonics **14**, 601-606 (2020).

[9]   A. Shalit, S. Ahmed, J. Savolainen, and P. Hamm, "Terahertz echoes reveal the inhomogeneity of aqueous salt solutions," Nature Chemistry **9**(2016).

[10]   K. K. Ng., S. M. Sze, *Physics of Semiconductor Devices,* Third ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007).

[11]   R. F. Pierret, *Semiconductor Device Fundamentals* (Addison-Wesley Publishing Company, Inc., 1996).

[12]   T. P. Chen, T. C. Lee, C. C. Ling, C. D. Beling, and S. Fung, "Current transport and its effect on the determination of the Schottky-barrier height in a typical system: Gold on silicon," Solid-State Electronics **36**, 949-954 (1993).

[13] X. Zhang and D. Joy, "Fabrication of Au/p-Si Schottky barrier for EBIC study," Microscopy Research and Technique **29**, no. 1: 47-53. (1994) doi:10.1002/jemt.1070290107.

[14] A. M. Cowley, "Titanium-silicon Schottky barrier diodes," Solid-State Electronics **13**, 403-414 (1970).

[15] R. H. Fowler, "The Analysis of Photoelectric Sensitivity Curves for Clean Metals at Various Temperatures," Physical Review **38**, 45-56 (1931).

[16] S. A. Maier, *Plasmonics: Fundamentals and Applications* (Springer, 2007).

[17] R. L. Olmon, B. Slovick, T. W. Johnson, D. Shelton, S.-H. Oh, G. D. Boreman, and M. B. Raschke, "Optical dielectric function of gold," Physical Review B **86**, 235147 (2012).

[18] M. A. Ordal, R. J. Bell, R. W. Alexander, L. A. Newquist, and M. R. Querry, "Optical properties of Al, Fe, Ti, Ta, W, and Mo at submillimeter wavelengths," Applied Optics **27**, 1203-1209 (1988).

[19] L. Vivien, L. Virot, D. Marris-Morini, J.-M. Hartmann, P. Crozat, E. Cassan, C. Baudot, F. Boeuf, and J.-M. Fédéli, "40Gbit/s germanium waveguide photodiode," in *Optical Fiber Communication Conference/National Fiber Optic Engineers Conference 2013*, OSA Technical Digest (online) (Optical Society of America, 2013), OM2J.3.

[20] H. Nyquist, "Thermal Agitation of Electric Charge in Conductors," Physical Review **32**, 110-113 (1928).

[21] H. J. Song *et al.*, "Uni-travelling-carrier photodiode module generating 300 GHz power greater than 1 mW," IEEE Microwave and Wireless Components Letters **22**, no. 7, pp. 363-365 (2012).

[22] C. C. Renaud *et al.*, "Antenna integrated THz uni-traveling carrier photodiodes," IEEE Journal of Selected Topics in Quantum Electronics **24**, no. 2, pp. 1-11 (2018).

[23] P. Latzel *et al.*, "Generation of mW Level in the 300-GHz band using resonant-cavity-enhanced unitraveling carrier photodiodes," IEEE Transactions on Terahertz Science and Technology **7**, no. 6, pp. 800-807 (2017).

[24] S. S. Mousavi, A. Stöhr, and P. Berini, "Plasmonic photodetector with terahertz electrical bandwidth," Applied Physics Letters **104**, no. 14, pp. 143112 (2014).

[25] A. Melikyan, L. Alloatti, A. Muslija, D. Hillerkuss, P. C. Schindler, J. Li, R. Palmer, D. Korn, S. Muehlbrandt, D. Van Thourhout, B. Chen, R. Dinu, M. Sommer, C. Koos, M. Kohl, W. Freude, and J. Leuthold, "High-speed plasmonic phase modulators," Nature Photonics **8**, 229-233 (2014).

[26] C. Haffner, W. Heni, Y. Fedoryshyn, J. Niegemann, A. Melikyan, D.L. Elder, B. Baeuerle, Y. Salamin, A. Josten, U. Koch, C. Hoessbacher, F. Ducry, L. Juchli, A. Emboras, D. Hillerkuss, M. Kohl, L. R. Dalton, C. Hafner, and J. Leuthold, "All-plasmonic Mach–Zehnder modulator enabling optical high-speed communication at the microscale," Nature Photonics **9**, 525-528 (2015).

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

[28] F. J. Ishi T, Makita K, Baba T, Ohashi K "Si Nano-Photodiode with a Surface Plasmon Antenna," Japanese Journal of Applied Physics **44**, L364 (2005).

[29] J. S. Helman and F. Sánchez-Sinencio, "Theory of Internal Photoemission," Physical Review B **7**, 3702-3706 (1973).

[30] M. W. Knight, H. Sobhani, P. Nordlander, and N. J. Halas, "Photodetection with Active Optical Antennas," Science **332**, 702-704 (2011).

[31] K.-T. Lin, H.-L. Chen, Y.-S. Lai, and C.-C. Yu, "Silicon-based broadband antenna for high responsivity and polarization-insensitive photodetection at telecommunication wavelengths," Nature Communications **5**, 3288, (2014).

[32] B. Desiatov, I. Goykhman, N. Mazurski, J. Shappir, J. B. Khurgin, and U. Levy, "Plasmonic enhanced silicon pyramids for internal photoemission Schottky detectors in the near-infrared regime," Optica **2**, 335-338 (2015).

[33] A. Akbari, R. N. Tait, and P. Berini, "Surface plasmon waveguide Schottky detector," Optics Express **18**, 8505-8514 (2010).

[34] M. Casalino, M. Iodice, L. Sirleto, I. Rendina, and G. Coppola, "Asymmetric MSM sub-bandgap all-silicon photodetector with low dark current," Optics Express **21**, 28072-28082 (2013).

[35] G. Q. L. Shiyang Zhu, and D. L. Kwong, "Low-Cost and High-Speed SOI Waveguide-Based Silicide Schottky-Barrier MSM Photodetectors for Broadband Optical Communications," IEEE Photonics Technology Letters **20**, 1396-1398 (2008).

[36] I. Goykhman, B. Desiatov, J. Khurgin, J. Shappir, and U. Levy, "Locally Oxidized Silicon Surface-Plasmon Schottky Detector for Telecom Regime," Nano Letters **11**, 2219-2224 (2011).

[37] C.-K. Tseng, W.-T. Chen, K.-H. Chen, H.-D. Liu, Y. Kang, N. Na, and M.-C. M. Lee, "A self-assembled microbonded germanium/silicon heterojunction photodiode for 25 Gb/s high-speed optical interconnects," Scientific Reports **3**, 3225 (2013).

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

[39] R. H. Fowler and L. Nordheim, "Electron Emission in Intense Electric Fields," Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **119**, 173-181 (1928).

[40] S. M. Sze, "Physics of Semiconductor Devices," in *Wiley-Interscience publication* (John Wiley & Sons, 1981).

[41] D. F. P. Pile and D. K. Gramotnev, "Adiabatic and nonadiabatic nanofocusing of plasmons by tapered gap plasmon waveguides," Applied Physics Letters **89**, 041111 (2006).

[42] H. Petek and S. Ogawa, "Femtosecond time-resolved two-photon photoemission studies of electron dynamics in metals," Progress in Surface Science **56**, 239-310 (1997).

[43] J.-M. Jeong, M. S. Oh, B. J. Kim, C.-H. Choi, B. Lee, C.-S. Lee, and S. G. Im, "Reliable Synthesis of Monodisperse Microparticles: Prevention of Oxygen Diffusion and Organic Solvents Using Conformal Polymeric Coating onto Poly(dimethylsiloxane) Micromold," Langmuir **29**, 3474-3481 (2013).

[44] R. W. Boyd, "Chapter 12 - Optically Induced Damage and Multiphoton Absorption," in *Nonlinear Optics (Third Edition)*, R. W. Boyd, ed. (Academic Press, Burlington, 2008), pp. 543-560.

[45] W. Freude, R. Schmogrow, B. Nebendahl, M. Winter, A. Josten, D. Hillerkuss, S. Koenig, J. Meyer, M. Dreschmann, M. Huebner, C. Koos, J. Becker, and J. Leuthold, "Quality metrics for optical signals: Eye diagram, Q-factor, OSNR, EVM and BER," in *14th International Conference on Transparent Optical Networks (ICTON)*, 2012, 1-4.

[46] Z. Xuezhe, P. Koka, M. O. McCracken, H. Schwetman, J. G. Mitchell, Y. Jin, R. Ho, K. Raj, and A. V. Krishnamoorthy, "Energy-Efficient Error Control for Tightly Coupled Systems Using Silicon Photonic Interconnects," IEEE/OSA Journal of Optical Communications and Networking **3**, A21-A31 (2011).

[47] T. Nagatsuma, G. Ducournau, and C. C. Renaud, "Advances in terahertz communications accelerated by photonics," Nature Photonics **10**, 371-379 (2016).

[48] T. Göbel, D. Stanze, B. Globisch, R. J. B. Dietz, H. Roehle, and M. Schell, "Telecom technology based continuous wave terahertz photomixing system with 105 decibel signal-to-noise ratio and 3.5 terahertz bandwidth," Optics Letters **38**, 4197-4199 (2013).

[49] S. Ummethala, T. Harter, K. Koehnle, Z. Li, S. Muehlbrandt, Y. Kutuvantavida, J. Kemal, P. Marin-Palomo, J. Schaefer, A. Tessmann, S. K. Garlapati, A. Bacher, L. Hahn, M. Walther, T. Zwick, S. Randel, W. Freude, and C. Koos, "THz-to-optical conversion in wireless communications using an ultra-broadband plasmonic modulator," Nature Photonics **13**, 519-524 (2019).

[50] Y. Salamin, P. Ma, B. Baeuerle, A. Emboras, Y. Fedoryshyn, W. Heni, B. Cheng, A. Josten, and J. Leuthold, "100 GHz Plasmonic Photodetector," ACS Photonics **5**, 3291-3297 (2018).

[51] M. L. Brongersma, "Plasmonic Photodetectors, Photovoltaics, and Hot-Electron Devices," Proceedings of the IEEE **104**, 2349-2361 (2016).

[52] A. P. Cracknell, Electron States and Fermi Surfaces of Elements · Ti: Datasheet from Landolt-Börnstein - Group III Condensed Matter · Volume 13C: "Electron States and Fermi Surfaces of Elements" in SpringerMaterials (https://doi.org/10.1007/10201682_57), Springer-Verlag Berlin Heidelberg.

[53] A. P. Cracknell, Electron States and Fermi Surfaces of Elements · Au: Datasheet from Landolt-Börnstein - Group III Condensed Matter · Volume 13C: "Electron States and Fermi Surfaces of Elements" in SpringerMaterials (https://doi.org/10.1007/10201682_7), Springer-Verlag Berlin Heidelberg.

[54] B. L. Smith and E. H. Rhoderick, "Schottky barriers on p-type silicon," Solid-State Electronics **14**, 71-75 (1971).

[55] S. Koenig, D. Lopez Diaz, J. Antes, F. Boes, R. Henneberger, A. Leuther, A. Tessmann, R. Schmogrow, D. Hillerkuss, R. Palmer, T. Zwick, C. Koos, W. Freude, O. Ambacher, J. Leuthold, and I. Kallfass, "Wireless sub-THz communication system with high data rate," Nature Photonics **7**, 977-981 (2013).

[56] J. Ma, N. J. Karl, S. Bretin, G. Ducournau, and D. M. Mittleman, "Frequency-division multiplexer and demultiplexer for terahertz wireless links," Nature Communications **8**, 729 (2017).

[57] J. Capmany and D. Novak, "Microwave photonics combines two worlds," Nature Photonics **1**, 319-330 (2007).

[58] K. Krügener *et al.*, "Terahertz meets sculptural and architectural art: evaluation and conservation of stone objects with T-ray technology," Scientific Reports **5**, 1482 (2015).

[59] M. Tonouchi, "Cutting-edge terahertz technology," Nature Photonics **1**, 97-105 (2007).

[60] S. Fan, Y. He, B. S. Ung, and E. Pickwell-MacPherson, "The growth of biomedical terahertz research," Journal of Physics D **47**, 374009 (2014).

[61] R. I. Stantchev, "Noninvasive, near-field terahertz imaging of hidden objects using a single-pixel detector," Science Advances **2**, e1600190 (2016).

[62] M. Naftaly, R. G. Clarke, D. A. Humphreys, and N. M. Ridler, "Metrology state-of-the-art and challenges in broadband phase-sensitive terahertz measurements," Proceedings of the IEEE **105**, no. 6, pp. 1151-1165 (2017).

[63] E. R. Brown, K. A. McIntosh, K. B. Nichols, and C. L. Dennis, "Photomixing up to 3.8 THz in low-temperature-grown GaAs," Applied Physics Letters **66**, 285-287 (1995).

[64] S. Preu, G. H. Döhler, S. Malzer, L. J. Wang, and A. C. Gossard, "Tunable, continuous-wave terahertz photomixer sources and applications," Journal of Applied Physics **109**, 061301 (2011).

[65] B. Sartorius, D. Stanze, T. Göbel, D. Schmidt, and M. Schell, "Continuous wave terahertz systems based on 1.5 μm telecom technologies," Journal of Infrared, Millimeter and Terahertz Waves **33**, 405-417 (2012).

[66] X. Yu *et al.*, "160 Gbit/s photonics wireless transmission in the 300–500 GHz band," APL Photonics **1**, 081301 (2016).

[67] R. Puerta *et al.*, "Single-carrier dual-polarization 328-Gb/s wireless transmission in a D-band millimeter wave 2×2 MU-MIMO radio-over-fiber system," Journal of Lightwave Technology **36,** no.2, pp. 587-593 (2018).

[68] Y. D. Hsieh *et al.*, "Terahertz frequency-domain spectroscopy of low-pressure acetonitrile gas by a photomixing terahertz synthesizer referenced to dual optical frequency combs," Journal of Infrared, Millimeter and Terahertz Waves **37**, 903-915 (2016).

[69] S. Hisatake, Y. Koda, R. Nakamura, N. Hamada, and T. Nagatsuma, "Terahertz balanced self-heterodyne spectrometer with SNR-limited phase-measurement sensitivity," Optics Express **23**, 26689-26695 (2015).

[70] T. Harter *et al.*, "Silicon-plasmonic photomixer for generation and homodyne reception of continuous-wave THz radiation," In *Conference on Lasers and Electro-Optics* SM4E.5 (OSA, 2016).

[71] T. Ishibashi, Y. Muramoto, T. Yoshimatsu, and H. Ito, "Unitraveling-carrier photodiodes for terahertz applications," IEEE Journal of Selected Topics in Quantum Electronics **20**, no. 6, pp. 79-88 (2014).

[72] A. Beling, X. Xie, and J. C. Campbell, "High-power, high-linearity photodiodes," Optica **3**, 328-338 (2016).

[73] C. W. Berry, "High power terahertz generation using 1550 nm plasmonic photomixers," Applied Physics Letters **105**, 011121 (2014).

[74] A. D. J. F. Olvera, H. Lu, A. C. Gossard, and S. Preu, "Continuous-wave 1550 nm operated terahertz system using ErAs:In(Al)GaAs photo-conductors with 52 dB dynamic range at 1 THz," Optics Express **25**, 29492-29500 (2017).

[75] D. Thomson *et al.* , "Roadmap on silicon photonics," Journal of Optics **18**, 073003 (2016).

[76] W. Bogaerts *et al.*, "Nanophotonic waveguides in silicon-on-insulator fabricated with CMOS technology," Journal of Lightwave Technology **23**, no. 1, pp 401-412 (2005).

[77] Vivien, L. & Pavesi, L. (eds) "Handbook of Silicon Photonics" (CRC Press, Boca Raton, FL, 2013).

[78] J. Pfeifle, L. Alloatti, W. Freude, J. Leuthold, and C. Koos, "Silicon–organic hybrid phase shifter based on a slot waveguide with a liquid-crystal cladding," Optics Express **20**, issue 14, pp. 15359-15376 (2012).

[79] N. C. Harris *et al.*, "Efficient, compact and low loss thermo-optic phase shifter in silicon," Optics Express **22**, 10487-10493 (2014).

[80] J. Y. Kim *et al.*, "Compact and stable THz vector spectroscopy using silicon photonics technology," Optics Express **22**, 7178-7185 (2014).

[81] G. T. Reed, G. Mashanovich, F. Y. Gardes, and D. J. Thomson, "Silicon optical modulators," Nature Photonics **4**, 518-526 (2010).

[82] C. Koos *et al.*, "Silicon-organic hybrid (SOH) and plasmonic-organic hybrid (POH) integration," Journal of Lightwave Technology **34**, no. 2, pp. 256-268 (2016).

[83] T. Hiraki *et al.*, "Heterogeneously integrated III–V/Si MOS capacitor Mach–Zehnder modulator," Nature Photonics **11**, 482-85 (2017).

[84] E. Timurdogan *et al.*, "An ultralow power athermal silicon modulator," Nature Communications **5**, 4008 (2014).

[85] M. R. Billah *et al.*, "Hybrid integration of silicon photonics circuits and InP lasers by photonic wire bonding," Optica **5**, 876-883 (2018).

[86] R. Quay *et al.*, "High-power microwaveGaN/AlGaN HEMTs and MMICs on SiC and silicon substrates for modern radio communication," Physica Status Solidi A, **215**: 1700655 (2018).

[87] G. V. Naik, V. M. Shalaev, and A. Boltasseva, "Alternative plasmonic materials: beyond gold and silver," Advanced Materials **25**: 3264-3294 (2013).

[88] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," IEEE Transactions on Electron Devices **22**, no. 11, pp. 1045-1047 (1975).

[89] J. Bowers and C. Burrus, "Ultrawide-band long-wavelength p–i–n photodetectors," Journal of Lightwave Technology **5**, no. 10, pp. 1339-1350 (1987).

[90] E. Rouvalis, M. J. Fice, C. C. Renaud, and A. J. Seeds, "Millimeter-wave optoelectronic mixers based on uni-traveling carrier photodiodes," IEEE Transactions on Microwave Theory and Techniques **60**, no. 3, pp. 686-691 (2012).

[91] S. Hisatake, J. Y. Kim, K. Ajito, and T. Nagatsuma, "Self-heterodyne spectrometer using uni-traveling-carrier photodiodes for terahertz-wave generators and optoelectronic mixers," Journal of Lightwave Technology **32**, no. 20, pp. 3683-3689 (2014).

[92] "International Roadmap For Devices And Systems 2021 Update: More Moore," (https://irds.ieee.org/images/files/pdf/2021/2021IRDS_MM.pdf, 2021).

[93] H. Esmaielpour, V. R. Whiteside, H. P. Piyathilaka, S. Vijeyaragunathan, B. Wang, E. Adcock-Smith, K. P. Roberts, T. D. Mishima, M. B. Santos, A. D. Bristow, and I. R. Sellers, "Enhanced hot electron lifetimes in quantum wells with inhibited phonon coupling," Scientific Reports **8**, 12473 (2018).

[94] D. K. T. Ng, W. Qian, L. Kim-Peng, P. Jing, T. Kun, L. Yicheng, L. Chee-Wei, and H. Seng-Tiong, "Effects of SiO2 hard masks on si nanophotonic waveguide loss for photonic device integration," IEEE Photonics Technology Letters **26**, 70-73 (2014).

F.   Glossary

List of abbreviations

AC	Alternating current
ADC	Analog-to-digital converter
Al	Aluminum
As	Arsenic
Au	Gold
BER	Bit-error ratio
BOX	Buried oxide
BW	bandwidth
C.w.	Continuous wave
CMOS	Complementary metal-oxide-semiconductor
DAC	Digital-to-analog converter
DC	Direct current
DFB	Distributed feedback laser
DOS	Density of states
EDFA	Erbium-doped fiber amplifier
eV	Electron Volt
FOM	Figure of merit
FPGA	Field-programmable arrays
Ga	Gallium
Ge	Germanium
HSQ	Hydrogen silsesquioxane

III–V semiconductor	Compound semiconductor from the third and the fifth main group elements
IMPATT	Impact ionization avalanche transit-time diode
In	Indium
IPE	Internal photoemission
LIA	Lock-in amplifier
LO	Local oscillator
MOS, MOSFET	Metal-Oxide-Semiconductor, MOS Field-effect transistor
MS	Metal-semiconductor
MSM	Metal-semiconductor-metal
MZM	Mach-Zehnder Modulator
NEP	Noise equivalent power
OOK	On-Off keying
PD	Photodetector
PIPED	Plasmonic internal photoemission detector
PM	Polarization-maintaining
PMMA	Polymethylmethacrylate
PRBS	Pseudo-random bit sequence
RF	Radio frequency
RIE	Reactive ion etching
Rx	Receiver
SEM	Scanning electron microscope
Si	Silicon
SiO_2	Silicon oxide
SMU	Source-measure unit
SNR	Signal-to-noise power ratio
SOI	Silicon-on-insulator

SPP	Surface plasmon polariton
TE	Transverse electric
Ti	Titanium
TIA	Transimpedance amplifier
TM	Transverse magnetic
T-wave	Terahertz wave
Tx	Transmitter
UTC	Uni-travelling-carrier
VNA	Vector network analyzer
WG	Waveguide

List of mathematical symbols

Greek symbols

μ	Carrier mobility
τ_{SC}	Scattering time
σ	Conductivity
$\varphi(x)$	Potential
ε	Permittivity
τ_{R}	Dielectric relaxation time
ϕ_{W}	Work function
$\phi_{\text{Au,e}}$	Barrier height (e.g. for electrons in Au)
χ_{Si}	Electron affinity
φ_{bi}	Built-in potential
ω	Angular frequency
δ	Penetration depth
α	Power attenuation constant

λ	Wavelength
Θ	Fiber orientation

Latin symbols

m^*	Effective mass of electron or hole
$\bar{n} = n - jn_i$	Complex refractive index
C	Capacitance
D_n, D_p	Diffusion constant (electron, hole)
e	Unit of charge
E_x	Electric field strength, in x-direction
f	Frequency
$f(W)$	Fermi-Dirac distribution function
g_n	Electron generation rate
$H(f)$	Transfer function
I_D	Dark current
I_P	Photocurrent
J	Electric current density
k	Propagation constant
L	Device length
n	Electron density
$N_C(W), N_V(W)$	Density of states (conduction band, valence band)
N_D, N_A	Dopant concentration (donator, acceptor)
p	Hole density
P, P_{opt}	Optical power
R	Resistance
r_n	Electron recombination rate

S	Sensitivity
t	Time
T	Temperature
U	Voltage
v, v_x	Carrier speed, in x-direction
W	Energy level
w	Semiconductor core width
W_F	Fermi energy level
W_G,	Band gap energy
x, y, z	Spatial coordinates

Danksagung

Diese Dissertationsschrift ist das Ergebnis meiner Promotion und entstand am Institut für Photonik und Quantenelektronik (IPQ) und am Institut für Mikrostrukturtechnik (IMT) am Karlsruher Institut für Technologie (KIT) im Zeitraum von 2012 bis 2017. Mit dieser Danksagung möchte ich meinen Mentoren und Weggefährten meinen Dank ausdrücken.

An erster Stelle gilt mein Dank meinem Doktorvater Prof. Christian Koos. Von seinen vielen Führungsqualitäten hat sein Streben nach Exzellenz auf mich wohl am stärksten gewirkt. Seine weitsichtigen Ziele motivieren dazu, auch vor großen Aufgaben nicht zurückzuschrecken, und mit seinen häufig unkonventionellen Lösungsideen hat er mir bei der Überwindung so mancher Herausforderung geholfen. Mein besonderer Dank gilt auch meinem Korreferenten Prof. Wolfgang Freude. Durch seine beeindruckende Auffassungsgabe, seinen ausgeprägten Wissensschatz und sein großes Interesse an unseren Forschungsergebnissen hat Prof. Freude ganz wesentliche Schwerpunkte meiner Arbeit gesetzt. Mir wird insbesondere der intensive Austausch zu wissenschaftlichen und vielen anderen Themen in Erinnerung bleiben, aus dem ich viel für meine Forschung und das Leben im Allgemeinen mitnehmen konnte. Ich danke meiner Korreferentin Prof. Jasmin Aghassi-Hagmann für ihre Begleitung während meiner Promotionsprüfung. Außerdem danke ich meinem Mentor Prof. Manfred Kohl, durch dessen Begleitung insbesondere bei allen Belangen der Nanofabrikation die Herstellung der PIPED erst ermöglicht wurde. Seine Erfahrung aus der Mikroaktorik und den zugehörigen Herstellungsprozessen war häufig ausschlaggebend um die PIPED Fabrikation zu optimieren. Ich danke außerdem Prof. Juerg Leuthold für seine Führung insbesondere zu Beginn meiner Promotion.

Ich danke den Kollegen Dr. Lothar Hahn, Peter Jakobs, Andreas Bacher, Alban Muslija, Dr. Martin Sommer und Ann-Kristin Nees für die tatkräftige Unterstützung bei der Herstellung der PIPED Bauteile. Ihr habt mit mir große Geduld bewiesen. Gleiches gilt für die Mitarbeiter des IPQ-Sekretariats und der Werkstatt, die mir stets bei allen organisatorischen Fragen, oder bei kniffligen Aufbauten, eine große Hilfe waren. Mein Dank gilt hier David Guder, Martin

Winkeler, Oswald Speck, Marco Hummel, Volker Bös, Bernadette Lehmann, Tatiana Gassmann, Andrea Riemensperger, Marlis Koch und zuletzt Alina Lauinger.

Ich danke den vielen ehemaligen Mitdoktoranden und meinen Studenten. Mir war die Zusammenarbeit mit euch stets eine große Freude und die Bewältigung jeder Herausforderung erschien möglich. Ihr seid eine herausragende Truppe.

Ich danke meiner geliebten Freundin Corinna, die ein Fels in meinem Leben ist. Zuletzt gilt mein Dank meinen geliebten Eltern Therese und Heinrich Mühlbrandt.

List of Publications

Journal Publications

[J1]   **Muehlbrandt, S.**, Harter, T., Fuellner, C., Ummethala, S., Wolf, S., Bacher, A., Hahn, L., Kohl, M., Freude, W., Koos, C., "**Field-effect silicon-plasmonic photodetector for coherent T-wave reception**," *Optics Express* **29**, 21586-21602 (2021) [DOI]

[J2]   Harter, T., Ummethala, S., Blaicher, M., **Muehlbrandt, S.**, Wolf, S., Weber, M., Adib, M. M. H., Kemal, J. N., Merboldt, M., Boes, F., Nellen, S., Tessmann, A., Walther, M., Globisch, B., Zwick, T., Freude, W., Randel, S., Koos, C., **"Wireless THz link with optoelectronic transmitter and receiver**," *Optica* **6**, 1063-1070 (2019) [DOI]

[J3]   Ummethala, S., Harter, T., Koehnle, K., Li, Z., **Muehlbrandt, S.**, Kutuvantavida, Y., Kemal, J. N., Marin-Palomo, P., Schaefer, J., Tessmann, A., Garlapati, S. K., Bacher, A., Hahn, L., Walther, M., Zwick, T., Randel, S., Freude, W., Koos, C., "**THz-to-optical conversion in wireless communications using an ultra-broadband plasmonic modulator**," *Nature Photonics* **13**, 519-524 (2019) [DOI]

[J4]   Harter, T., Muehlbrandt, S., Ummethala, S., Schmid, A., Nellen, S., Hahn, L., Freude, W., Koos, C., "**Silicon–plasmonic integrated circuits for terahertz signal generation and coherent detection**," *Nature Photonics* **12**, 625-633 (2018) [DOI]

[J5]   Paasch-Colberg, T., Kruchinin, S. Yu., Saglam, O., Kapser, S., Cabrini, S., **Muehlbrandt, S.**, Reichert, J., Barth, J. V., Ernstorfer, R., Kienberger, R., Yakovlev, V. S., Karpowicz, N., Schiffrin, A., "**Sub-cycle optical control of current in a semiconductor: from the multiphoton to the tunneling regime**," *Optica* **3**, 1358-1361 (2016) [DOI]

[J6] **Muehlbrandt, S.**, Melikyan, A., Harter, T., Koehnle, K., Muslija, A., Vincze, P., Wolf, S., Jakobs, P. -J., Fedoryshyn, Y., Freude, W., Leuthold, J., Koos, C., Kohl, M., "**Silicon-plasmonic internal-photoemission detector for 40 Gbit/s data reception**," *Optica* **3**, 741-747 (2016) [DOI]

[J7] Koos, C., Leuthold, J., Freude, W., Kohl, M., Dalton, L. R., Bogaerts, W., Giesecke, A. L., Lauermann, M., Melikyan, A., Koeber, S., Wolf, S., Weimann, C., **Muehlbrandt, S.**, Koehnle, K., Pfeifle, J., Hartmann, W., Kutuvantavida, Y., Ummethala, S., Palmer, R., Korn, D., Alloatti, L., Schindler, P. C., Elder, D. L., Wahlbrink, T., Bolten, J., "**Silicon-Organic Hybrid (SOH) and Plasmonic-Organic Hybrid (POH) Integration**," *Journal of Lightwave Technology* **34**, 01-13 (2016) [DOI] more, (invited)

[J8] Melikyan, A., Koehnle, K., Lauermann, M., Palmer, R., Koeber, S., **Muehlbrandt, S.**, Schindler, P. C., Elder, D. L., Wolf, S., Heni, W., Haffner, C., Fedoryshyn, Y., Hillerkuss, D., Sommer, M., Dalton, L. R., Thourhout, D. V., Freude, W., Kohl, M., Leuthold, J., Koos, C., "**Plasmonic-organic hybrid (POH) modulators for OOK and BPSK signaling at 40 Gbit/s**," *Optics Express* **23**, 9938-9946 (2015) [DOI]

[J9] Melikyan, A., Alloatti, L., Muslija, A., Hillerkuss, D., Schindler, P. C., Li, J., Palmer, R., Korn, D., **Muehlbrandt, S.**, Thourhout, D. V., Chen, B., Dinu, R., Sommer, M., Koos, C., Kohl, M., Freude, W., Leuthold, J., "**High-speed plasmonic phase modulators**," *Nature Photonics* **8**, 229-233 (2014) [DOI]

[J10] Leuthold, J., Hoessbacher, C., **Muehlbrandt, S.**, Melikyan, A., Kohl, M., Koos, C., Freude, W., Dolores-Calzadilla, V., Smit, M., Suarez, I., Martinez-Pastor, J. P., Fitrakis, E. P., Tomkos, I., "**Plasmonic Communications: Light on a Wire**," *Optics and Photonics News* **24**, 28-35 (2013) [DOI]

[J11] Schiffrin, A., Paasch-Colberg, T., Karpowicz, N., Apalkov, V., Gerster, D., **Mühlbrandt, S.**, Korbman, M., Reichert, J., Schultze, M., Holzner, S., V. Barth, J., Kienberger, R., Ernstorfer, R., Yakovlev, V.S., Stockman, M. I., Krausz, F., "**Optical-field-induced current in dielectrics**,", *Nature* **493**, 70-74 (2013)

Conference Publications

[C1] Koos, C., Randel, S., Freude, W., Zwick, T., Scheytt, J. C., Witzens, J., Walther, M., Harter, T., Ummethala, S., Kieninger, C., Zwickel, H., Marin-Palomo, P., **Muehlbrandt, S.**, Fuellner, C., Schaefer, J., Gudyriev, S., Zazzi, A., Mueller, J., Tessmann, A., "**Photonic-Electronic Ultra-Broadband Signal Processing: Concepts, Devices, and Applications**," *2020 IEEE BiCMOS and Compound Semiconductor Integrated Circuits and Technology Symposium (BCICTS'20), Monterey (CA) USA, pp. 1-3, Nov. 16-19* (2020) [DOI]

[C2] Freude, W., Ummethala, S., Harter, T., Koehnle, K., Li, Z., **Muehlbrandt, S.**, Kutuvantavida, Y., Kemal, J. N., Marin-Palomo, P., Schaefer, J., Tessmann, A., Kumar Garlapati, S., Bacher, A., Hahn, L., Walther, M., Zwick, T., Randel, S., Koos, C., "**Wireless THz-to-optical conversion with an electro-optic plasmonic modulator**," *Light: Science & Applications, Light Conference (Light Conference'19), Changchun, China, July 16-18* (2019) more (invited)

[C3] Freude, W., Harter, T., **Muehlbrandt, S.**, Ummethala, S., Nellen, S., Hahn, L., Randel, S., Koos, C., "**Silicon photonics for coherent terahertz generation and detection**," *21th Intern. Conf. on Transparent Optical Networks (ICTON'19), Angers, France, pp. 1-4, July 9-13,* (2019) [DOI] (invited)

[C4]   Freude, W., Harter, T., Ummethala, S., **Muehlbrandt, S.**, Blaicher, M., Wolf, S., Weber, M., Boes, F., Massler, H., Tessmann, A., Kutuvantavida, Y., Kemal, J. N., Nellen, S., Hahn, L., Globisch, B., Walther, M., Zwick, T., Randel, S., Koos, C., "**Wireless terahertz communications**," *Conf. on Lasers and Electro-Optics (CLEO/Europe-EQEC 2019), Munich, Germany, paper cc_5_1, June 23-27* (2019) [DOI]

[C5]   Freude, W., Harter, T., Ummethala, S., **Muehlbrandt, S**., Blaicher, M., Wolf, S., Weber, M., Boes, F., Massler, H., Tessmann, A., Kutuvantavida, Y., Kemal, J. N., Nellen, S., Hahn, L., Goebel, T., Walther, M., Zwick, T., Randel, S., Koos, C., "**Terahertz communications meets photonics**," *14th Intern. Conf. on Fiber Optics & Photonics (Photonics'18), Indian Institute of Technology Delhi (IIT Delhi), Delhi, India, December 12–15* (2018) (invited)

[C6]   Ummethala, S., Harter, T., Koehnle, K., Li, Z., **Muehlbrandt, S.,** Kutuvantavida, Y., Kemal, J. N., Schaefer, J., Massler, H., Tessmann, A., Kumar Garlapati, S., Bacher, A., Hahn, L., Walther, M., Zwick, T., Randel, S., Freude, W., Koos, C., "**Wireless Transmission at 0.3 THz Using Direct THz-to-Optical Conversion at the Receiver**," *44th European Conference on Optical Communication (ECOC'18), Rome, Italy, September 23-27* , paper We4H.3 (2018) [DOI] (invited)

[C7]   Harter, T., Ummethala, S., **Muehlbrandt, S**., Blaicher, M., Koehnle, K., Adib, M. M. H., Weber, M., Wolf, S., Kutuvantavida, Y., Kemal, J. N., Nellen, S., Hahn, L., Tessmann, A., Walther, M., Zwick, T., Randel, S., Freude, W., Koos, C., "**Wireless terahertz communications using optoelectronic techniques**," *Progress in Electromagnetics Research Symposium 2018 (PIERS 2018), Toyama, Japan, August 1-4* (2018) (invited)

[C8]   Koos, C., Randel, S., Freude, W., Dalton, L. R., Wolf, S., Kieninger, C., Kutuvantavida, Y., Lauermann, M., Elder, D. L., **Muehlbrandt, S.**, Zwickel, H., Melikyan, A., Harter, T., Ummethala, S., Billah, M. R., Blaicher, M., Dietrich, P.-I., Hoose, T., "**Hybrid Photonic Integration and Plasmonic Devices: New Perspectives for High-Speed Communications and Ultra-Fast Signal Processing**," *Pacific Rim Conference on Lasers and Electro-Optics (CLEO-PR 2018), Hong Kong, July 29 – August 3* , pp. 1-2 (2018) [DOI] more (invited)

[C9]   Freude, W., Harter, T., Ummethala, S., **Muehlbrandt, S.**, Blaicher, M., Wolf, S., Weber, M., Boes, F., Massler, H., Tessmann, A., Kutuvantavida, Y., Kemal, J. N., Nellen, S., Hahn, L., Goebel, T., Walther, M., Zwick, T., Randel, S., Koos, C., "**Wireless THz communications using optoelectronic techniques**," *Light: Science & Applications, Light Conference (Light Conference'18), Changchun, China, July 16-18* (2018) (invited)

[C10]   Ummethala, S., Harter, T., Koehnle, K., **Muehlbrandt, S.**, Kutuvantavida, Y., Kemal, J. N., Schaefer, J., Massler, H., Tessmann, A., Kumar Garlapati, S., Bacher, A., Hahn, L., Walther, M., Zwick, T., Randel, S., Freude, W., Koos, C., "**Terahertz-to-Optical Conversion Using a Plasmonic Modulator**," *Conf. on Lasers and Electro-Optics (CLEO'18), San Jose (CA), USA, May 13–18* , paper STu3D.4 (2018) [DOI] more

[C11]   Harter, T., Adib, M. M. H., Wolf, S., **Muehlbrandt, S.**, Weber, M., Blaicher, M., Boes, F., Massler, H., Tessmann, A., Nellen, S., Goebel, T., Giesekus, J., Walther, M., Zwick, T., Freude, W., Randel, S., Koos, C., "**Wireless Multi-Subcarrier THz Communications Using Mixing in a Photoconductor for Coherent Reception**," *IEEE Photonics Conference (IPC2017), Orlando, Florida, USA, 1-5 October* , paper MA4.2 (2017) [DOI]

[C12] Koos, C., Freude, W., Guber, A. E., Schimmel, T., Lauermann, M., Schneider, S., Weimann, C., **Muehlbrandt, S.**, Harter, T., Wondimu, S. F., Wienhold, T., von der Ecken, S., Dietrich, P.-I., Goering, G., "**Photonic Integration for Metrology and Sensing**," *OSA Integrated Photonics Research, Silicon, and Nano- Photonics (IPR), New Orleans, Louisiana, USA, 24 - 27 July,* paper ITh1A.1. (2017) [DOI] (invited)

[C13] Freude, W., **Muehlbrandt, S.**, Harter, T., Melikyan, A., Koehnle, K., Muslija, A., Vincze, P., Wolf, S., Jakobs, P. -J., Fedoryshyn, Y., Leuthold, J., Kohl, M., Zwick, T., Randel, S., Koos, C., "**PIPED: A silicon-plasmonic high-speed photodetector**," *19th Intern. Conf. on Transparent Optical Networks (ICTON'17), Girona, Spain, July 2-6*, paper Mo.D4.1 (2017) [DOI] (invited)

[C14] Harter, T., Weber, M., **Muehlbrandt, S.**, Wolf, S., Kemal, J. N., Boes, F., Nellen, S., Goebel, T., Giesekus, J., Zwick, T., Randel, S., Freude, W., Koos, C., "**Wireless THz Communications Using Optoelectronic Techniques for Signal Generation and Coherent Reception**," *Conf. on Lasers and Electro-Optics (CLEO'17), San Jose (CA), USA, May 14 – 19*, paper SM3J.2. (2017) [DOI] more

[C15] Koos, C., Freude, W., Leuthold, J., Dalton, L. R., Wolf, S., **Muehlbrandt, S.**, Melikyan, A., Zwickel, H., Kutuvantavida, Y., Harter, T., Lauermann, M., Elder, D. L., "**Nanophotonic modulators and photodetectors using silicon photonic and plasmonic device concepts**," *SPIE Physics and Simulation of Optoelectronic Devices XXV (OPTO-SPIE'17), San Francisco (CA), USA, Jan. 28 - Feb. 2* , paper 10098-6 (2017) [DOI] (invited)

[C16] Melikyan, A., Koehnle, K., Lauermann, M., Palmer, R., Koeber, S., **Muehlbrandt, S.**, Schindler, P. C., Elder, D. L., Wolf, S., Sommer, M., Dalton, L. R., Thourhout, D. V., Freude, W., Kohl, M., Leuthold, J., Koos, C., "**Plasmonic-organic hybrid (POH) modulators**," *21st OptoElectronics and Communications Conference (OECC) held jointly with International Conference on Photonics in Switching (PS'2016), Niigata, Japan, 3-7 July* , paper ThE3-1 (2016) (invited)

[C17] Harter, T., **Muehlbrandt, S.**, Ummethala, S., Schmid, A., Bacher, A., Hahn, L., Kohl, M., Freude, W., Koos, C., "**Silicon-Plasmonic Photomixer for Generation and Homodyne Reception of Continuous-Wave THz Radiation**," *Conf. on Lasers and Electro-Optics (CLEO'16), San Jose (CA), USA, June 5 – 10*, paper SM4E.5. (2016) [DOI]

[C18] Koos, C., Freude, W., Leuthold, J., Kohl, M., Dalton, L. R., Bogaerts, W., Lauermann, M., Wolf, S., Weimann, C., Melikyan, A., Lindenmann, N., Billah, M. R., **Muehlbrandt, S.**, Koeber, S., Palmer, R., Koehnle, K., Alloatti, L., Elder, D. L., Giesecke, A. L., Wahlbrink, T., "**Silicon-organic hybrid (SOH) integration and photonic multi-chip systems: Extending the capabilities of the silicon photonic platform**," *28th Annual Conf. of the IEEE Photonics Society (IPC'15), Reston (VA), USA, Oct. 4–8* , paper MF2.1 (2015) [DOI] (invited)

[C19] Koos, C., Freude, W., Leuthold, J., Kohl, M., Dalton, L. R., Bogaerts, W., Lauermann, M., Melikyan, A., Wolf, S., Weimann, C., **Muehlbrandt, S.**, Pfeifle, J., Koeber, S., Palmer, R., Schindler, P. C., Elder, D. L., "**Silicon-organic (SOH) and plasmonic-organic (POH) hybrid integration: Extending the capabilities of silicon photonics and plasmonics**," *20th OptoElectronics and Communications Conference (OECC'15), Shanghai, China, June 28 – July 2* , paper JTuE.34 (2015) [DOI] (invited)

[C20] Melikyan, A., Koehnle, K., Lauermann, M., Palmer, R., Koeber, S., **Muehlbrandt, S.**, Schindler, P. C., Elder, D. L., Wolf, S., Heni, W., Haffner, C., Fedoryshyn, Y., Hillerkuss, D., Sommer, M., Dalton, L. R., Van Thourhout, D., Freude, W., Kohl, M., Leuthold, J., Koos, C., "**Plasmonic-organic hybrid (POH) modulators for OOK and BPSK signaling at 40 Gbit/s**," *Conf. on Lasers and Electro-Optics (CLEO'15), San Jose (CA), USA, May 10–15* , paper SM1I.1. *Optical Society of America (OSA)* (2015) [DOI]

[C21] **Muehlbrandt, S.**, Melikyan, A., Koehnle, K., Harter, T., Muslija, A., Vincze, P., Wolf, S., Jakobs, P. -J., Fedoryshyn, Y., Freude, W., Leuthold, J., Koos, C., Kohl, M., "**Plasmonic Internal Photoemission Detectors with Responsivities above 0.12 A/W**," *Conf. on Lasers and Electro-Optics (CLEO'15), San Jose (CA), USA, May 10–15* , paper FTh3E.3 *Optical Society of America (OSA)* (2015) [DOI]

[C22] Koos, C., Leuthold, J., Freude, W., Kohl, M., Dalton, L. R., Bogaerts, W., Giesecke, A. L., Lauermann, M., Melikyan, A., Koeber, S., Wolf, S., Weimann, C., **Muehlbrandt, S.**, Koehnle, K., Pfeifle, J., Palmer, R., Alloatti, L., Elder, D. L., Wahlbrink, T., Bolten, J., "**Silicon-Organic Hybrid (SOH) and Plasmonic-Organic Hybrid (POH) Integration**," *Optical Fiber Communication Conference (OFC'15), Los Angeles (CA), USA, 22 – 26 March* , paper Tu2A.1 (2015) [DOI] (invited)

[C23] Koehnle, K., Lauermann, M., Bacher, A., Nees, M. -K., **Muehlbrandt, S.**, Palmer, R., Muslija, A., Jakobs, P. -J., Freude, W., Koos, C., "**Silicon-Organic Hybrid Phase Modulator Based on Sub-Wavelength Grating Waveguides**," *Micro and Nano Engineering (MNE2014), Lausanne, Switzerland, September 22-26*, Paper 8127, Session: D2L-B (2014)

[C24] Melikyan, A., Alloatti, L., Muslija, A., Hillerkuss, D., Schindler, P. C., Li, J., Palmer, R., Korn, D., Lindenmann, N., **Muehlbrandt, S.**, Walheim, S., Vincze, P., Leufke, P. M., Ulrich, S., Ye, J., Thourhout, D. V., Chen, B., Dinu, R., Sommer, M., Hahn, H., Schimmel, T., Koos, C., Kohl, M., Freude, W., Leuthold, J., "**High-speed Plasmonic Modulators**," *Integrated Photonics Research, Silicon and Nanophotonics (IPR'14), San Diego, California United States, July 13-17* , paper IT2A.6 *Optical Society of America (OSA)* (2014) [DOI] (invited)

[C25] **Muehlbrandt, S.**, Muslija, A., Koehnle, K., Melikyan, A., Leuthold, J., Kohl, M., "**Fabrication of Ultra-Compact Plasmonic Waveguide Photo Diodes**," *Micro and Nano Engineering (MNE'2014), Lausanne, Switzerland*, paper 8274 (2014)

[C26] Melikyan, A., Alloatti, L., Muslija, A., Hillerkuss, D., Schindler, P. C., Li, J., Palmer, R., Korn, D., **Muehlbrandt, S.**, Thourhout, D. V., Chen, B., Dinu, R., Sommer, M., Koos, C., Kohl, M., Freude, W., Leuthold, J., "**Surface Plasmon Polariton High-Speed Modulator**," *Conf. on Lasers and Electro-Optics (CLEO'13), San Jose (CA), USA, June 9–14* , paper CTh5D.2 *Optical Society of America* (2013) [DOI] (postdeadline)

[C27] Melikyan, A., Sommer, M., Muslija, A., Kohl, M., **Muehlbrandt, S.**, Mishra, A. K., Calzadilla, V., Justo, Y., Martinez-Pastor, J. P., Tomkos, I., Scandurra, A., Van Thourhout, D., Hens, Z., Smit, M., Freude, W., Koos, C., Leuthold, J., "**Chip-to-chip plasmonic interconnects and the activities of EU project NAVOLCHI**," *14th International Conference on Transparent Optical Networks (ICTON), 2012, July 02-05*, pp. 1-3, University of Warwick, Coventry, UK, July (2012) [DOI] (invited)

Karlsruhe Series in Photonics & Communications, Vol. 39
Edited by Profs. C. Koos, W. Freude and S. Randel

Karlsruhe Institute of Technology (KIT)
Institute of Photonics and Quantum Electronics (IPQ)
Germany

—IPQ—
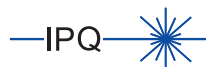
High-speed photodetectors are core elements of optoelectronic systems. The seamless cointegration of photonic and electronic components is currently impeded by vastly different boundary conditions imposed by the underlying technologies. Whereas modern field-effect transistors can be fabricated with feature sizes approaching the single-digit nanometer range, typical dimensions of integrated photodetectors amount to several micrometers. Furthermore, while modern CMOS electronics are fabricated on the mature silicon platform, photodetectors are typically fabricated in germanium or on dedicated III–V substrates. This book introduces and experimentally demonstrates a novel approach to exploit internal photoemission (IPE) for photodetection in silicon. We call the resulting device PIPED (plasmonic IPE detector). PIPED represents the first detector based on IPE that operates at technically viable performance figures. It combines a compact footprint and a high bandwidth with monolithic cointegration in silicon photonic circuits. PIPED open an attractive route towards optoelectronic generation and detection of waveforms in the sub-THz and THz range, so called T-waves.

About the Author

Sascha Mühlbrandt was born 1985 in Stuttgart, Germany. He received the Physics diploma (Dipl.-Phys.) in 2011 from Technical University of Munich (TUM). In 2023, he was awarded the Dr.-Ing. (Ph.D.) degree in Electrical Engineering and Information Technology from Karlsruhe Institute of Technology (KIT). His research focuses on high-speed signal processing.

ISBN 978-3-7315-1367-4

9 783731 513674