Engineering Novel Semiconductors for Advanced Infrared Sensors

Seokhyeong Lee

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2023

Reading Committee:
Mo Li, Chair
Arka Majumdar
Sajjad Moazzeni

Program Authorized to Offer Degree:
Electrical and Computer Engineering

University of Washington

Abstract

Engineering Novel Semiconductors for Advanced Infrared Sensors

Seokhyeong Lee

Chair of the Supervisory Committee:
Mo Li
Department of Electrical and Computer Engineering

Infrared (IR) technologies have boundless applications spanning from fundamental science, industrial engineering to bio-medical science, climate and environment science. The technology has been developed with quantum physics, solid state physics, and semiconductor engineering, with deeper understanding of interaction between light and matter especially via charged particles (also quasi-particles). Heterogeneous integration of different semiconductor materials adds more functionalities to optoelectronic devices, where properties of optical material can be modulated by engineering the nearby materials. In two-dimensional materials, the effects from perturbation from the near environment can be large enough to change optical properties of the materials. In addition, the vertical confinement of two-dimensional material exhibits enhanced light-matter interaction with quantum confinement effect and reduced electron screening, becoming great candidates for optoelectronic devices. Black Phosphorus (bP) is thermodynamically stable two-dimensional material with direct bandgap spanning broad infrared regime. The main two chapters utilize bP for IR active material. First, the optical properties, absorption coefficient can be modulated by electric field and electrostatic doping using Field Effect Transistor (FET) geometry, governed by physical phenomena of Burstein-Moss effect and Quantum Confined Franz Keldysh effect. The modulation of absorption coefficient of bP and its integrability on Silicon chip enables on-chip optical modulator. Second chapter where bP is utilized offers additional functionality, non-volatile programmability by engineering dielectric layer of FET geometry. The non-volatile various values of responsivities of bP IR sensors, when combined

in an array, serve as in-sensor computing processor where simple neural network computation is possible in sensor using physical/analog computing, to demonstrate edge detection and hand written digit recognition with 92 % accuracy. Finally, III-V compound semiconductor system is considered and engineered to capture 1-2 μm of broad near-IR photons with exceptional sensitivity and photoelectric gain. Here, heterojunction phototransistor (HPT) is conceived and optimized to not only detect IR light but also drive an LED to extend its application to light-weight Night Vision Intensifier, exhibiting unprecedented IR photon to Red photon upconversion.

Acknowledgements

The journey for Ph.D. has taught me a lot, so much as I consider myself blessed with full of grateful things. I still remember the rejoice I felt when I decided to join professor Mo Li's group, after a year of wondering with research topic in University of Minnesota. My personal interests in light and solid state physics perfectly matches the expertise of prof. Mo Li, and his broad and cutting-edge research intrigued me to dive into the photonic devices research. Less than a year later I joined the group, the whole laboratory moved to University of Washington, which was another fortunate thing in my life, because it helped me engage every detail of the laboratory from inventory of equipment and tools to installation of a new laboratory. Nevertheless, research takes time, efforts, and patience with more failures than success in experiments and proof of ideas. It would have been impossible to finish the journey without the help of my advisor, colleagues, collaborators, friends, and family. I would like to express my sincere gratitude to them.

First of all, I genuinely appreciate professor Mo Li for offering the opportunity to explore the field of photonics and nanotechnology with great expertise so that I can pursue the Ph.D. degree. Through the last six years including moving from Minnesota, prof. Mo Li has taught me a lot, not only for science and technology but also for growing as an independent researcher, with his great patience and encouragement. Especially, his instruction on expressing and delivering scientific ideas in writing and presentation was one of the most valuable teachings I have ever had.

I also want to thank my senior colleagues: Dr. Huan Li, Dr. Li He, Dr. Ruoming Peng, Dr. Che Chen, Dr. Qiyu Liu, Dr. Bingzhao Li, Dr. Changming Wu, who has been with me from University of Minnesota when I just joined the group and gave me lots of instructions, broadening ideas, and struggling together with difficult problems. I am proud of them for their impacts on research and industry. Also I want to thank my current group members: I-Tung Chen, Adina Ripin, Qixuan Lin, Haoqing Deng, Shucheng Fang, Hanna

Boyer, Yue Ye. It was great help and encouragement to have talented colleagues to work with and inspire each other. We have been through the COVID shut-down and tools and equipment shut-downs to make progress on research with encouragement and idea exchanges.

I want to extend my gratitude to my collaborators and friend researchers from other groups: Prof. Arka Majumdar, Prof. Sajjad Moazeni, Prof. Alexandra Velian, Dr. Johannes E. Fröch, Dr. Zhuoran Feng, Dr. Abhi Saxena, Dr. Quentin Tanguy, Dr. Minho Choi, Dr. Cheng Chang, Arnab Manna, Michael Riehs, Chris Munley, Hannah Rarick, and Rui Chen in UW, Prof. Lawrence Minjoo Lee and Dr. David Alexander Montealegre in UIUC, Prof. Fengnian Xia and Yihao Song in Yale. Their companionship, guidance, and encouragement were unforgettable experience, helping my growth as an individual researcher, with sharing ideas and resources for better environment of research.

Also, I want to express my gratitude to all the technicians, engineers, and administrators in research facilities providing me with hands-on experience and resources for research: Molecular Analysis Facility (MAF), the National Nanotechnology Coordinated Infrastructure site at the University of Washington(UW NNCI), Washington Nanofabrication Facility (WNF), and University of Washington (UW) Molecular Engineering Materials Center (MEM-C).

Lastly, I want to thank my parents, Sangmok Lee and Sangae Choi, and my brother, Sumoon Lee for their endless support and loving for more than six years of journey.

Last but not the least, I thank Ghayoung Kim, my Fiance, for all the support in the darkest times and patiently listening to my complains when experiments and publication didn't work out. With all the love and support, the completion of degree was possible.

DEDICATION

To my Family

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Light's impact on humanity extends beyond its role as an energy source for daily sustenance. Study of light has resulted in the evolution from the geomteric-ray theory to the wave theory and finally to the quantum theory. Its profound influence extends to the very foundation of quantum mechanics and modern physics, marking the inception of modern civilization actively harnessing electromagnetic force.

Its interaction with matter, or light-matter interaction has led to numerous scientific achievements, governing many (almost all) of the chemical and physical phenomena observed in daily life and beyond. From the boundless reaches of the cosmos to the microscopic landscapes of quantum realms, the interaction between light and matter stands as a cornerstone of scientific exploration and technological innovation.

Through meticulous analysis, experimental insights, and theoretical frameworks, this work aspires to illuminate the intricate mechanisms and phenomena underpinning light-matter interaction and semiconductor engineering. It seeks to unravel the complexities of energy/charge transfer, quantum effects, and circuit analysis, opening avenues for revolutionary advancements in fields such as materials science, telecommunications, and IR imaging technologies.

1.1 Motivation

In the vast spectrum of electromagnetic radiation, the realm of the infrared (IR) holds a wealth of potential, offering a bridge between the visible and microwave regions. The pursuit of harnessing IR radiation has led to a burgeoning field of photonics and semiconductor physics, unveiling a myriad of possibilities for groundbreaking technological advancements. Table 1.1 summarizes the brief history of infrared technologies developed with quantum theory and semiconductor technologies, beginning from its discovery in 1800s by William Herschel.

At the heart of this exploration lies the remarkable properties of semiconductor materials tailored to manipulate and harness infrared light. Semiconductor physics forms the bedrock upon which innovative IR photonics technologies are built. This thesis endeavors to dissect the underlying principles governing the behavior of elementary particles such as electrons, phonons, and photons within these materials, paving the way for a deeper understanding and unlocking the potential for novel applications in IR photonic devices. This work aims to unravel the mysteries of semiconductor-based infrared photonics, pushing the boundaries of what is conceivable in sensing, communications, computing, medical diagnostics, and beyond. Studying charged carriers and their dynamic behaviors in materials is crucial for optimizing material utilization and achieving desired functionalities efficiently.

1.2 Infrared Photodetectors

The foundation of modern infrared (IR) technology traces back to significant discoveries and advancements spanning over a century. In 1887, Hertz's pioneering work revealed that ultraviolet irradiation of a conductor emitted negatively charged particles, marking the inception of photoemission. Subsequent experiments in the late 19th century demonstrated the use of visible radiation with alkali metal electrodes, expanding the boundaries of light-matter interaction [4, 5]. Ferdinand Braun's 1874 discovery of rectifying properties at semiconductor-metal contacts, using naturally occurring lead sulphide crystals, laid the groundwork for semiconductor technology. Further exploration by Jagadis Chandra Bose, around 1901, led to the patenting of a semiconductor rectifier, propelling advancements in radio signal detection [6, 7]. Theodore W. Case's 1917 development of the first IR photoconductor using thallium and sulphur (Tl_2S) marked an early

Year	Event
1800	Discovery of the existence of thermal radiation in the invisible beyond the red by W. HERSCHEL
1821	Discovery of the thermoelectric effects using an antimony-copper pair by T.J. SEEBECK
1833	Thermopile consisting of 10 in-line Sb-Bi thermal pairs by L. NOBILI and M. MELLONI
1834	Discovery of the PELTIER effect on a current-fed pair of two different conductors by J.C. PELTIER
1835	Formulation of the hypothesis that light and electromagnetic radiation are of the same nature by A.M. AMPERE
1840	Discovery of the three atmospheric windows by J. HERSCHEL (son of W. HERSCHEL)
1859	Relationship between absorption and emission by G. KIRCHHOFF
1864	Theory of electromagnetic radiation by J.C. MAXWELL
1874	Ferdinand Braun's discovery of rectifying properties at semiconductor-metal contact
1879	Empirical relationship between radiation intensity and temperature of a blackbody by J. STEFAN
1884	Thermodynamic derivation of the STEFAN law by L. BOLTZMANN
1887	Observation of photoelectric effect in ultraviiolet by H. HERTZ
1894	Derivation of the wavelength relation of blackbody radiation by J.W. RAYEIGH and W. WIEN
1900	Discovery of quantum properties of light by M. PLANCK
1901	J.C. BOSE obtained a patent for PbS point-contact rectifiers
1905	A. EINSTEIN established the theory of photoelectricity
1911	The first television image tube using cathode ray tubes by R. ROSLING
1917	T. Case developed the first IR photoconductor using thallium and sulphur
1923	W. SCHOTTKY established the theory of dry rectifiers
1928	F. Bloch developed the theory of electrons in lattices
1930	IR direction finders based on PbS quantum detectors in the wavelength range 1.5-3.0 μm for military applications (GUDDEN, GORLICH and KUTSCHER)
1935	EINSTEIN, PODOLSKY, and ROSEN proposed a thought experiment regarding paradoxical quantum entanglement
1939	First IR display unit in the United States (Sniperscope, Snooperscope)
1941	R.S. OHL observed the photovoltaic effect shown by a p-n junction in a silicon
1942	G. EASTMAN (Kodak) offered the first film sensitive to the infrared
1954	First imaging cameras based on thermopiles and on bolomters
1955	Mass production start of IR detectors in the US (PbS and PbTe detectors)
1957	Discovery of HgCdTe ternary alloy as infrared detector material by W.D. LAWSON, S. NELSON, and A.S. YOUNG
1965	Mass production start of IR cameras for civil applicatrions in Sweden (single-element sensors with mechanical scanner)
1970	Discovery of charge-couple device (CCD) by W.S. BOYLE and G.E. SMITH
1975	First In-bump hybrid infrared focal plane array
1977	Discovery of the type-II InAs/GaSb superlattices by G.A. SAI-HALASZ, R. TSU, and L. ESAKI
1980	First demonstration of SWIR InGaAsP detector by J.C. CAMPBELL, A.G. DENTAI, T.P. LEE, and C.A. BURRUS)
1985	Development and mass prouction of cameras fitted with Schottky diode FPAs
1990	Development of quantum well infrared photoconductor (QWIP) hybrid systems
2004	Demonstration of two-dimensional Material by A. GEIM at al.
2014	Two dimensional black phosphorus has regained attention by researchers

Table 1.1: Discovery of Infrared and development of Semiconductor physics and technology (modified with permission from Ref.[2] and Ref.[3]). Copyright 2002. Elsevier Science

milestone in IR detection. Despite initial promise, these detectors faced challenges such as instability in the presence of light, high noise, and lack of reproducibility, leading to discontinued efforts in 1918 [8]. The concept of the electro-optical converter emerged in 1928, culminating in the creation of a working tube with a photocathode and a fluorescent screen by Holst et al. in 1934. Japanese scientists Asao and Suzuki's 1940 method for enhancing silver's sensitivity in the S-1 photocathode extended its usefulness into the near-infrared region [9, 10]. During World War II, the Radio Corporation of America (RCA) developed the 1P25 image converter, a significant advancement in near-IR (NIR) technology. Post-war, various photocathodes were developed, enabling detection across different spectral regions [7]. Efforts in the early 1930s by Edgar W. Kutzscher et al. discovered lead sulphide's photoconductive properties, leading to the development of sensitive PbS films in Germany. This work, done under secrecy, resulted in the fabrication of the most sensitive German detectors, deployed in various wartime applications [11, 12, 13]. In the post-war era, the trajectory of detector technology, initially led by German research, continued in the USA, Great Britain, and the Soviet Union. This knowledge dissemination post-1945 significantly influenced subsequent advancements globally [14]. Although initially developed for military use, low-cost IR imagers now find applications across various nonmilitary domains, including industrial monitoring, night vision, law enforcement, and rescue operations [15, 16, 17]. The evolution of IR technology, stemming from a sequence of pioneering discoveries and advancements, has transformed it from a wartime asset into a versatile and indispensable tool serving diverse modern applications.

1.2.1 Classification of infrared detectors

Reviewing the evolution of infrared (IR) detector technology post-World War II, numerous materials have undergone extensive investigation. Following Norton's proposition [18], a fundamental observation emerges: "An array of physical phenomena within the approximate energy range of 0.1 to 1 electron-volt (eV) will be suggested for application in IR detectors." Among these effects are: thermoelectric power (thermocouples)[19], change in electrical conductivity (bolometers)[20, 21], gas expansion (Golay cell)[22, 23], pyroelectricity (pyroelectric detectors) [24, 25], photon drag, Josephson effect (Josephson junctions, Superconduction Quantum Interference Detector (SQUID))[26, 27], internal emission (PtSi Schottky barriers)[28], fundamental absorption (intrinsic photodetectors)[29, 30, 31], low dimensional solids (superlattice (SL), quan-

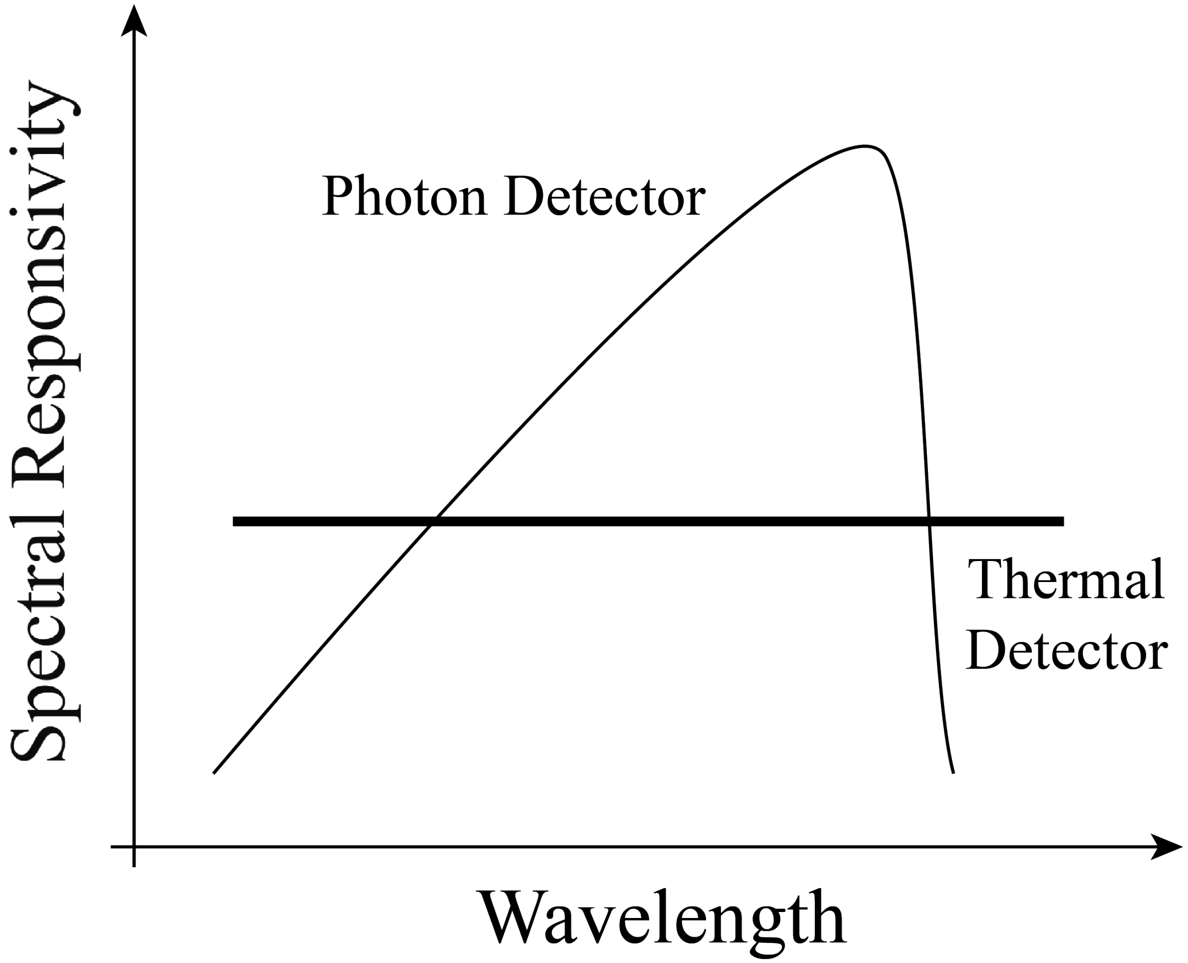


Figure 1.1: Infrared detectors can be categorized mainly in two: Photon detectors and Thermal detectors

tume well (QW), quantum dot (QD), Van-der-Waals semiconductor detectors), different type of phase transitions, etc. The World War II era marked the inception of contemporary IR detector technology. The notable progress in employing infrared technology for addressing remote sensing challenges owes its success to the effective evolution of high-functioning infrared detectors spanning the past sixty years. The convergence of photon IR technology, semiconductor material science, photolithography techniques tailored for integrated circuits, and the urgency of military readiness during the Cold War era have collectively driven remarkable strides in IR capabilities within a relatively brief span in the last century [32]. The majority of optical detectors can be classified in two broad categories: photon detectors (also called quantum detectors) and thermal detectors.

1.2.2 Photon detectors

In photon detectors, Infrared radiation is absorbed within the material by interaction with electrons either bound to lattice atoms or to impurity atoms, or with free electrons. The alteration in electronic energy distribution leads to the observed electrical output signal. Photon detectors display a specific wavelength-dependent response concerning the incident radiation power. They demonstrate excellent signal-to-noise ratios and rapid responsivities. However, achieving these qualities in infrared photon detectors demands cryogenic cooling to avert the generation of charge carriers due to thermal excitation. Without cooling, thermal transitions interfere with optical ones, resulting in heightened noise levels in non-cooled devices. The spectral current responsivity of photon detectors is equal to

$$R_i = \frac{\lambda\eta}{hc} qg, \tag{1.1}$$

where λ is the wavelength, h is the Planck's constant, c is the velocity of light, q is the electron charge, and g is the photoelectric current gain. The current that flows through the contacts of the device is noisy due to the statistical nature of the generation and recombination processes - fluctuation of optical generation, thermal generation, and radiative and nonradiative recombintation rates. Assuming that the current gain for the photocurrent and the noise current are the same, the noise current is

$$I_n^2 = 2q^2g^2(G_{op} + G_{th} + R)\Delta f, \tag{1.2}$$

where G_{op} is the optical generation rate, G_{th} is the thermal generation rate, R is the resulting recombination rate, and Δf is the frequency band. Jone [33] found that for many detectors the noise equivalent power (NEP) is proportional to the square root of the detector signal that is proportional to the detector area, A_d. The normalized detectivity, or specific detectivity D^* is defined by as:

$$D^* = \frac{(A_d)^{1/2}}{NEP}. \tag{1.3}$$

Detectivity, D^*, is the prevalent parameter to characterize normalized signal-to-noise performance of detectors and can be also defined as

$$D^* = \frac{R_i(A_d \Delta f)^{1/2}}{I_n}. \tag{1.4}$$

The importance of D^* is that this figure of merit permits comparison of detectors of the same type, but having different areas. Either a spectral or blackbody D^* can be defined in terms of corresponding type of NEP. At equilibrium, the generation and recombination rates are equal. In this case,

$$D^* = \frac{\lambda \eta}{2hc(Gt)^{1/2}}. \tag{1.5}$$

Background radiation usually is the main source of noise in an IR detector. Assuming no contribution due to recombination,

$$I_n^2 = 2\Phi_B A_d \eta q^2 g^2 \Delta f, \tag{1.6}$$

where Φ_B is the background photon flux density. Therefore, at the background limited performance conditions (BLIP performance),

$$D^*_{BLIP} = \frac{\lambda}{hc}(\frac{\eta}{\Phi_B})^{1/2}. \tag{1.7}$$

Once background-limited performance is reached, quantum efficiency, η, is the only detector parameter that can influence a detector's performance. Depending on the nature of the interaction, the class of photon detectors is further sub-divided into different types. The most important are: intrinsic detectors, extrinsic detectors, photoemissive (Schotky barriers). Different types of detectors can be referred to reference *Infrared Detectors* [32]. Figure 1.1 shows spectral responsivity curves for a number of commercially available IR detectors.

1.2.3 Thermal detector

The second class of detectors is composed of thermal detectors. Thermal detectors represent a class of detectors that operate based on the principle of converting incident radiation into a change in material temperature. These detectors harness the resulting alteration in a physical property to generate an electrical output. Unlike photon-based detectors, the signal in thermal detectors is independent of the photonic nature of the incoming radiation. Instead, it hinges on changes in radiation power or the rate of change, making the signal intensity the pivotal factor rather than the spectral content of the radiation. Typically, a thermal detector is constructed with the sensor suspended on legs that connect it to a heat sink. This setup ensures efficient dissipation of heat, preventing thermal drift and maintaining stability in the detector's response. Due to its design, the detector's response is not limited by wavelength; it captures incident radiation indiscriminately across various wavelengths. This broad spectral response is facilitated by the ability to absorb radiation using a black surface coating. Among the various types of thermal detectors, three primary approaches have garnered significant attention and utility within the realm of infrared technology: Bolometers, Pyroelectric Detectors, and Thermoelectric detectors.

The versatility and effectiveness of thermal detectors, particularly bolometers, pyroelectric, and thermoelectric detectors, have made them invaluable in the field of infrared technology. Their ability to operate across a wide spectral range and their sensitivity to changes in radiation power or rate change make them indispensable tools in various applications, including thermal imaging, spectroscopy, remote sensing, and industrial monitoring.

1.3 Dissertation Outline

The work comprising this dissertation is arranged as follows:

- Chapter 1 introduces important feats in Infrared (IR) technologies and its development with Semiconductor and Solid-state-physics.
- Chapter 2 discusses the theory of semiconductor materials for optoelectronics and delves into two material systems with great potentials used in the following chapters.
- Chapter 3 introduces Black Phosphorus integration on Silicon photonic waveguides for optical detectors and modulators
- Chapter 4 presents the experimental results of programmable photo-transistors using black phosphorus for theri unique optoelectronic functionalities and the application in novel computing
- Chapter 5 discusses III-V material based-phototransistor with unprecedentedly high gain and its application to light-weighted night vision intensifiers
- Chapter 6 gives a summary of the work presented in this disseratation along with a discussion of future perspectives.

Chapter 2

Infrared Optoelectronic Materials

2.1 Direct Low-Bandgap Materials

With establishment of quantum physics and growing understanding of solid-state physics, semiconductor industry has experienced exponential growth with variety of purified materials. Silicon has achieved notable success through its meticulous carrier doping and advanced nano-fabrication techniques, enabling the creation of CMOS transistors at the atomic-molecular scale for accelerated electronic computation. Despite these accomplishments, silicon's inherent indirect bandgap limits its application in optical active devices, such as on-chip lasers, efficient LEDs, and modulators. Furthermore, the bandgap confines silicon to the visible and a small portion of the Near Infrared Range (up to 950 nm), excluding it from the richer information and opportunities present in the infrared spectrum.

Near Infrared (NIR) regime, spanning from 800 nm to 2 μm, holds significant value in applications such as optical communication, subcutaneous and ophthalmological tomography, biosensing, optical computing, and Light Detection and Ranging (LiDAR) [34, 35, 36, 37]. Beyond this, Mid-Infrared (MIR) range, extending from 2 μm to 12 μm, boasts characteristic absorption lines that serve as unique fingerprints for chemical and bio-chemical molecules, making it well-suited for chemical detection and biomedical characterization [38, 39, 40].

The bandgap of a material plays a pivotal role in determining its optical and electronic properties, thereby influencing the performance of optical devices. Materials with a direct bandgap are particularly favored for optical transitions, and the bandgap energy must align with the wavelengths of interest. Black Phosphorus, a 2-Dimensional material, exhibits a direct bandgap with tunable energy ranging from 0.3 eV to 1.6 eV, dependent on its thickness. This characteristic renders Black Phosphorus highly versatile for achieving optimal optoelectronic functionality in the Near Infrared (NIR) and Mid-Infrared (MIR) ranges [41]. III-V materials, derived from the third and fifth columns of the periodic table, boast a rich history of study spanning decades. This extensive research has successfully identified various semiconductor compositions within this class, enabling the development of LEDs, lasers, and photodetectors with direct bandgaps across a range of bandgap energies.

This chapter centers on the exploration of two prominent optical materials: black phosphorus, as a low-bandgap 2-D material, and InGaAsP, belonging to the III-V material class. A brief but comprehensive analysis of the optical and electronic properties of these materials, derived from their atomic structures, will be undertaken. Additionally, we will delve into strategic engineering approaches and methodologies aimed at realizing specific application functionalities based on the unique characteristics of each material. Furthermore, the investigation will extend to the integration of these materials in emerging technologies, highlighting their potential impact on advancing optical devices and electronic applications in the following chapters.

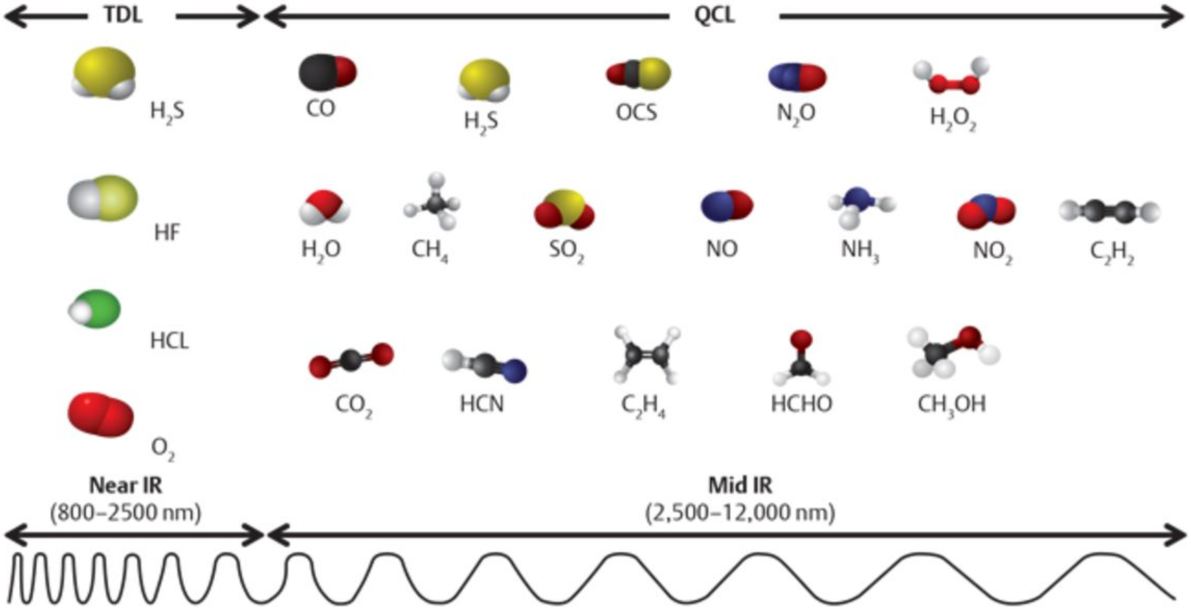


Figure 2.1: Molecular Absorption lines spanning Near and Mid Infrared wavelength regime. Characteristic spectral lines corresponding to vibration modes are considered to be "finger print" of certain molecules [42]

2.2 Two-Dimensional Semiconductor

Atomically thin, two-dimensional materials, so-called 2D materials, offer vertical confinement of electronic, phononic, and photonic system, being a platform for strong interaction between quantum particles. The realization of graphene in 2004 with an unexpectedly easy method [3] expedited the progress and more discoveries on 2D materials. Taking the monolayer transition metal dichalcogenides (TMDC) as an example, the reduced dielectric environment in vertical direction contributes to a stronger excitonic response, the optical absorption and photoluminescence (PL) are 2 orders of magnitude larger than the direct band-to-band transitions [43]. Different 2D materials vary in their electronic bandgap from insulating hexagonal Boron Nitride (hBN), through semiconducting transition metal dichalcogenides (TMDC) and black Phosphorus (bP), to semimetal materials such as graphene [44]. These materials can cover very broad electromagnetic spectral range from microwaves to ultraviolet as shown in Fig 2.2.

2.2.1 2D Materials with Light-Matter Interaction

The quantum confinement of 2D materials and consequent enhanced interaction with light serves a platform for various nanophotonic engineering and research with novel phenomena and quasi-particles, such as excitons [45, 46, 47]. Not only those unrepresented physical phenomena, can 2D materials' light-matter interaction extend its application to advanced photodetectors. Note that the family of 2D materials encompass broad range of material with different physical properties, including bandgap, carrier mobility, and thermal conductivity, etc. This allows for the creation of a completely autonomous next generation of optoelectronics that relies solely on 2D materials, which can be chosen and tailored based on the unique requirements of each application, such as sensitivity, speed, or wavelength spectrum coverage.

2.2.2 Emergence of Graphene Optoelectronics and Problems

Graphene, monolayer carbon atoms forming a hexagonal honeycomb lattice, is a gapless Dirac semimetal material with ultrahigh mobility[3, 48, 49], being an appealing material for broadband photodetection from terahertz (THz) to ultraviolet (UV)[50, 51, 52] and ultrafast technologies [53, 54]. Graphene can also serve as a transparent electrode whose work function is tunable due to linear dispersion of the Dirac electrons near the K point. These superior properties of graphene drew huge attention in optoelectronic fields. Graphene's

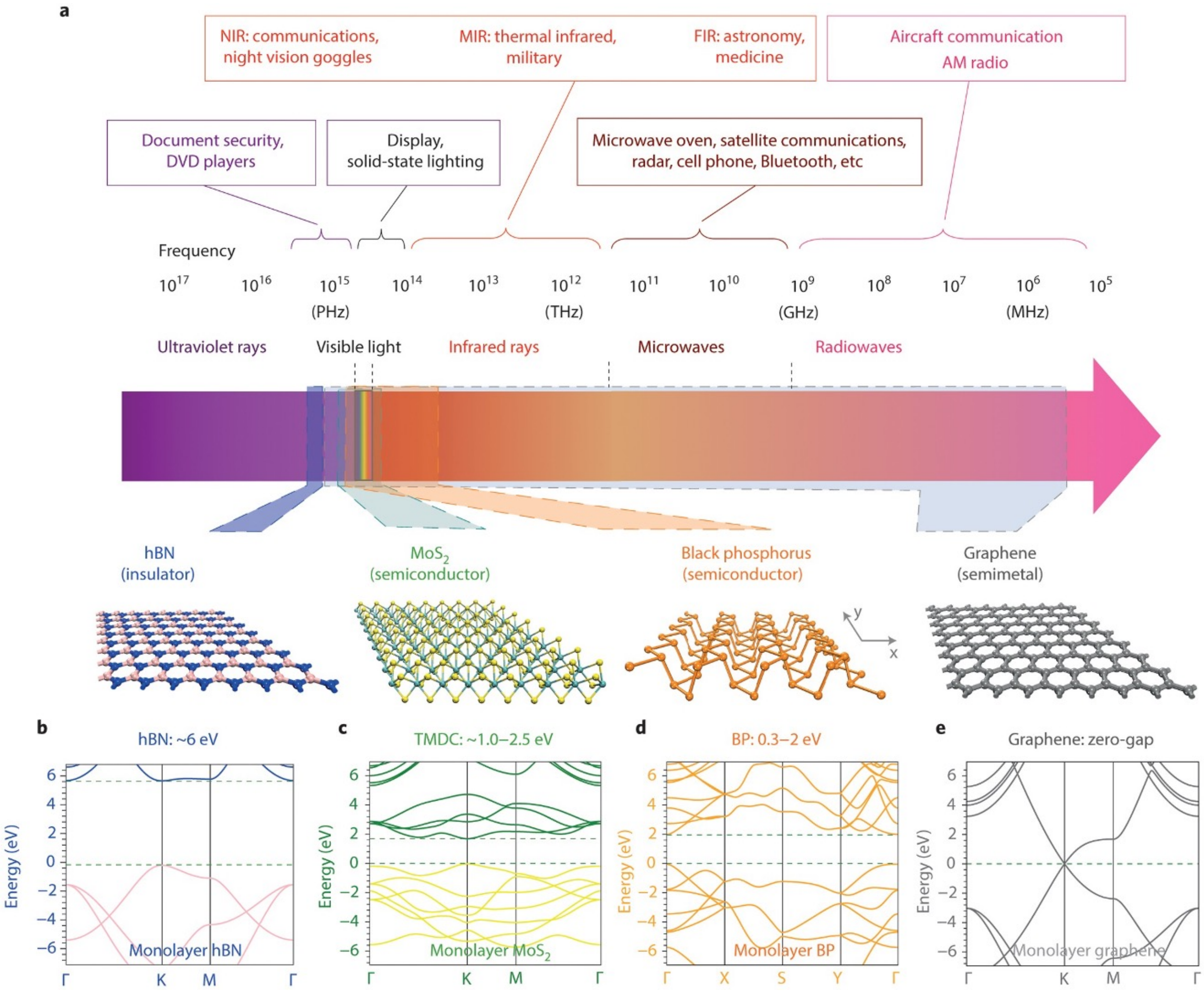


Figure 2.2: Bandgap and corresponding optical frequencies of various 2-dimensional materials. Reproduced with permission from [44]. Copyrights 2014 Authors. Nature Publishing Group

high electrical conductivity combined with transparency makes it an ideal candidate for transparent electrodes in optoelectronic devices like touchscreens, organic light-emitting diodes (OLEDs), and solar cells. It outperforms traditional materials like indium tin oxide (ITO) in terms of flexibility and conductivity [55, 56, 57]. Graphene's unique band structure allows for light emission in the terahertz range, potentially revolutionizing communication and imaging technologies [58, 59]. Among many other optoelectronic devices, graphene exhibits a broad spectral response, from ultraviolet to infrared wavelengths, enabling its use in photodetectors. Its high carrier mobility allows for quick and efficient photoresponse [60, 61].

Despite its promising potential, the integration of graphene into optoelectronic devices faces several challenges. Graphene's lack of an intrinsic bandgap restricts its use in some optoelectronic applications, like light-emitting diodes (LEDs), where a bandgap is essential for emitting light efficiently. Also, the bandgapless nature of graphene results in large dark current when it is used for photodetector, limiting detectivity or sensitivity of the photodetector. While graphene's properties are exceptional, tailoring them for specific optoelectronic applications often requires modifications that might alter its intrinsic properties.

We will introduce and change our focus to black phosphorus with direct bandgap around 0.3 eV (tunable with thickness) for infrared optoelectronic applications. Small but finite bandgap plays an important role in IR photodetector and other optoelectronic applications.

2.2.3 Black Phosphorus and its Optoelectronic Properties

Black Phosphorus (bP) has regained scientific attention when a few research teams reintroduced it from the perspective of a layered thin film material [41] at the beginning of 2014. Black Phosphorus, the most thermodynamically stable allotrope of the phosphorus element, is emerging as a promising semiconductor with a moderate band gap for nanoelectronics and nanophotonics applications [60]. Unlike in group IV elemetal layered materials, such as graphene, silicene, or germanene, each phosphorus atom has five outer shell electrons, with puckered orthorhombic crystalline structure which belongs to D_{2h}^{18} point group (2.2d) and has reduced symmetry compared to its group IV counterparts. This results in unique anisotropic electronic and optical properties. Also, black phosphorus possesses high carrier mobility, especially in its few-layer or monolayer forms. This property is advantageous for electronic devices as it facilitates the rapid movement of charge carriers, enabling high-speed electronic applications [62]. Combined with high mobility and strong light-matter interaction, bP is suitable for applications such as photodetectors and photovoltaic devices. Its high light absorption coefficient allows efficient conversion of photons into electrical signals. Figure 2.3 shows the band structure of bulk black phosphorus and its bandgap change with the number of layers. Consequently, the electronic properties of black phosphorus, such as bandgap and carrier mobility, are thickness-dependent. As the number of layers decreases, the bandgap increases, offering a means to tailor its electronic behavior for specific applications.

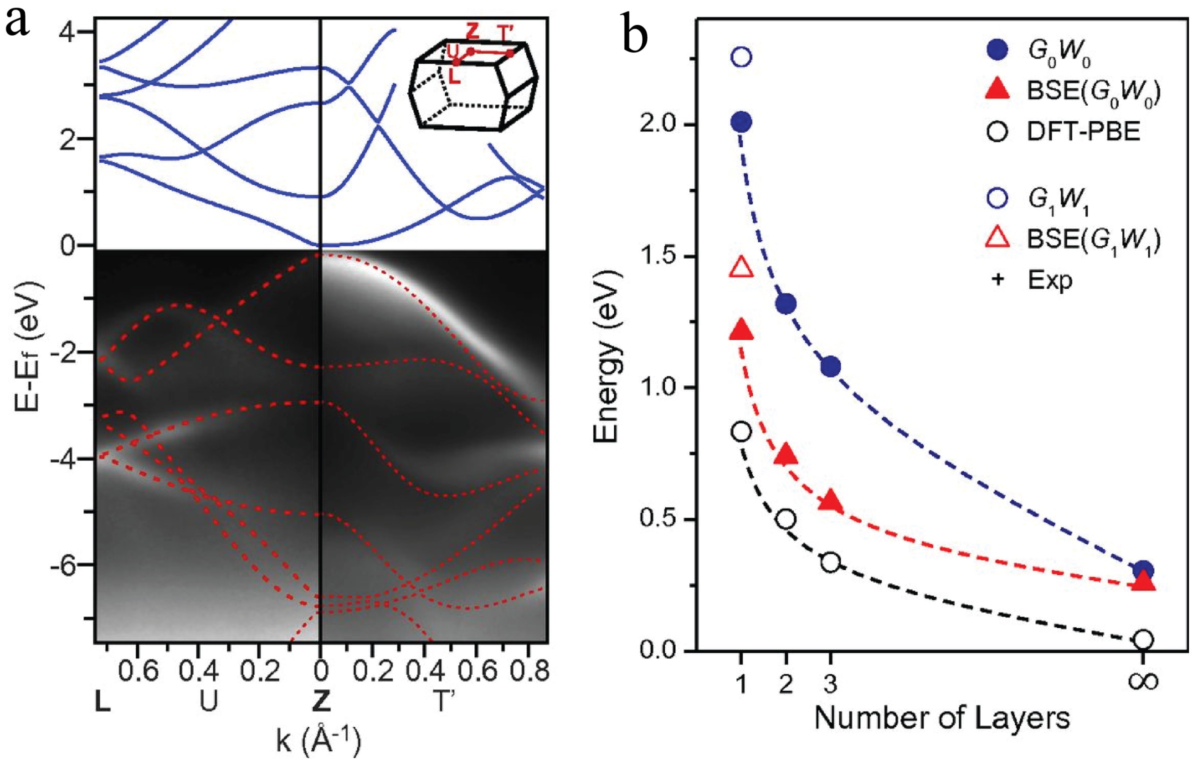


Figure 2.3: (a) Band Structure of bulk black Phosphorus mapped out by ARPES measurements. A band gap around 0.3 eV is clearly observed. Superimposed on top are the calculated bands of the bulk crystal. Blue solid and red dashed lines denote empty and filled bands, respectively. The direction of the mapping are along U (L-Z) and T', as indicated in the first Brillion zone shown in Inset. E_f is the Fermi Energy. (b) The evolution of the band gap calculated by different methods, and the energy of the optical absorption peak according to the stacking layer number of few-layer bP. Reproduced with permission from [63, 64]

2.2.4 Black Phosphorus for Infrared Technologies

Mid-Infrared (mid-IR) technology has drawn significant attention recently. Due to the increasing demands for the wide branch of applications, including free-space communication, beam steering, bio-chemical detection, and gas sensing [65, 38], varied novel photonic devices have been proposed where the high-speed detectors and modulators are highly desirable. For the free-space communications, the atmosphere offers several transparent windows in the mid-IR regime, mainly at 3-5 µm and 8-14 µm. Moreover, the development of mid-IR photonics will also benefit the research in chemistry and biology. Varied chemical molecules and gases have strong vibration absorption in the mid-IR regime, which can be used as the identifiers for the molecule identification and gas sensing. Although the mid-IR photonics has become crucial for the development of optical communication and chemical sensing, the commercial devices for the mid-IR regime still rely on the traditional bulk materials including InAsSb and HgCdTe which are usually high cost and unable to integrate with the Si photonic platform. As an alternative, the 2D materials such as bP and graphene have already been demonstrated for the photodetection and electro-optical modulation in the mid-IR regime. Especially, black phosphorus (bP) has the unique property that its bandstructure varies with the thickness due to its strong interlayer coupling. In results, the bandgap of the bP is layer dependent ranging from around 0.3 eV in the bulk to approximately 1.5-2.0 eV in monolayers (the corresponding optical wavelength range is 0.62-4.3 µm). Figure 2.4 illustrates various optoelectronic application using bP. BP in Field effect transistor (FET) geometry serves a tunable IR photodetector with high responsivity [60] and optical modulators [66]. When combined with other 2D materials, bP-based heterostructure can extend its functionality; BP/InSe heterostructure forms PN-junction, exhibiting avalanche breakdown, serving as mid-IR avalanche photodetector [67]. Also, light emitting devices are also crucial application in optoelectronics. BP as an direct bandgap material with 0.3 eV, plays a role as an LED whose wavelength peaks at 3.6 μm, as shown in Fig. 2.4. These versatile applications of bP suggest the importance of research and engineering based on the material. With engineering the dielectric structure of field effect transistor, bP phototransistor with programmability would be fabricated, characterized, and utilized to demonstrate smart image sensor.

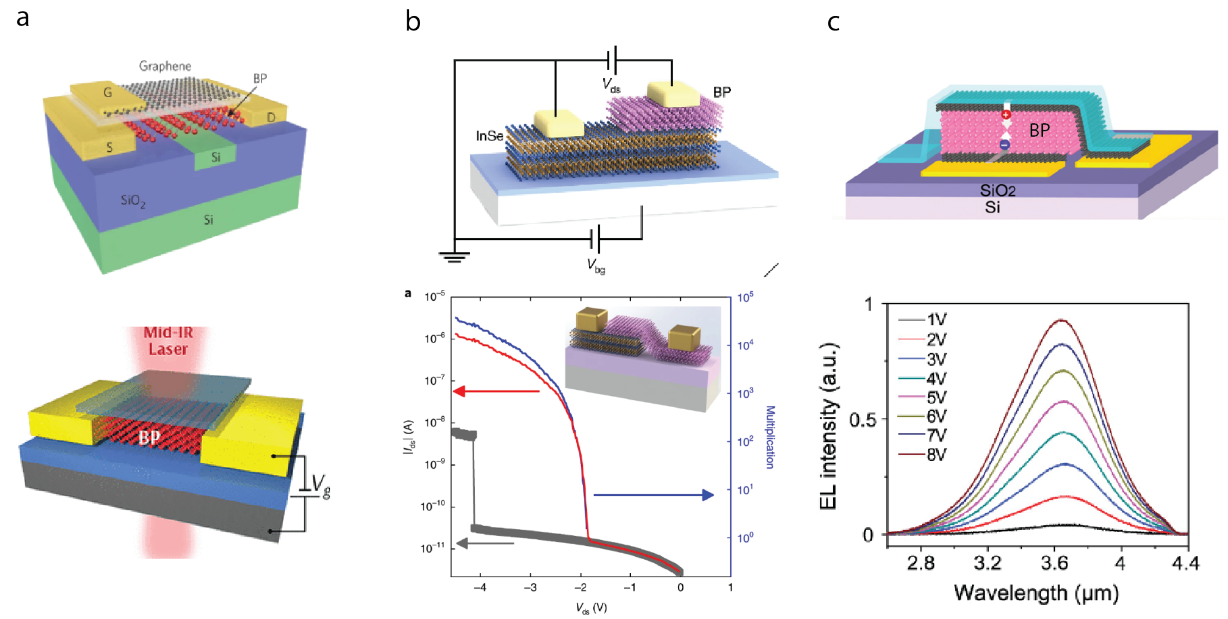


Figure 2.4: (a) Black Phosphorus in FET geometry used for IR photodetector (top) and modulator (bottom)[60, 66] (b) Heterostructure consisting of bP and InSe can exhibit avalanche effect and serve as an avalanche photodetector. Bottom plot shows the avalanche photocurrent with 4 μm light and corresponding multiplication factor [**?**]. (c) BP being direct low-bandgap material emits light in mid-IR regime by fabricated as an LED [68]. The figures are rearranged with the permission from [60, 66, 67, 68]. Copyright. Nature Publishing Group. ACS publication

2.3 III-V semiconductor

III-V compound semiconductors have played a crucial role in the development of optoelectronic devices, contributing significantly to fields such as photonics, telecommunications, and optoelectronic integrated circuits. III-V compound semiconductors are materials composed of elements from Group III (e.g., Al, Ga, In) and Group V (e.g., N, P, As) of the periodic table. Common examples include Gallium Arsenide (GaAs), Indium Phosphide (InP), and Gallium Nitride (GaN). One key advantage of III-V semiconductors is the ability to engineer the bandgap by selecting different combinations of elements. This allows the design of materials with specific optical and electronic properties. Figure 2.5 shows the lattice constants and bandgap of III-V semiconductors depending on its composition, enables lattice matched material growth with minimal strains with bandgap engineering for various purposes.

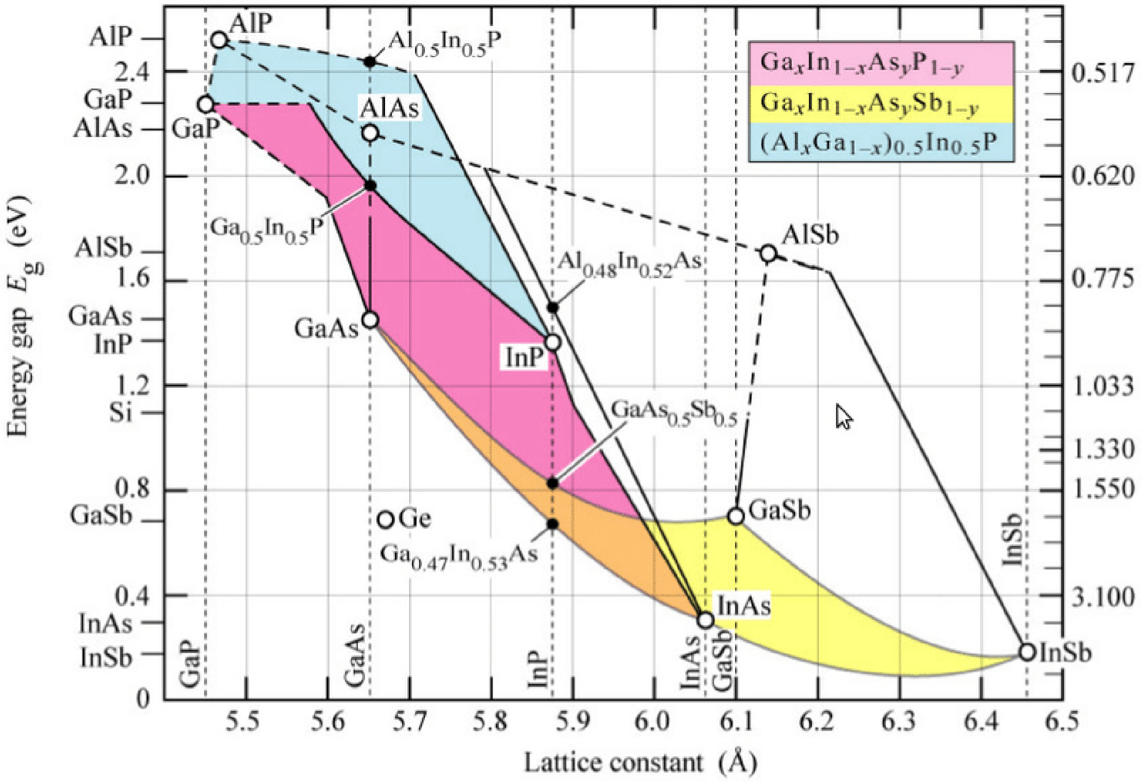


Figure 2.5: Bandgap energy vs. lattice constant at room temperature for various III-V semiconductors and their alloys. Circles represent binaries, solid and dashed lines ternaries, and areas between them quaternaries. Reproduced with permission from [69]

2.3.1 Compound Semiconductors

As mentioned above, the bandgap engineering is pivotal in optoelectronic applications. Compound Semiconductor is a term encompassing broad families of semiconductor materials formed by combining elements from different groups of the periodic table, creating compounds with unique electronic properties. One prominent group within compound semiconductors is the III-V semiconductor family, composed of elements from group III (e.g., gallium, indium) and group V (e.g., nitrogen, phosphorus) and II-VI semiconductor family, composed of elements from group II and group VI (e.g., Zinc Sulfide, Cadmium Telluride, Mercury Cadmium Telluride). Depending on bandgap engineered by selecting the combination of materials in proper amounts, the wide range of photodetection with optimized efficiency is possible. Figure 2.6 shows different applications depending on the bandgap. S. Nakamura [70] first demsontrated blue LED using wide bandgap compound semiconductor baed on InGaN. III-V materials' bandgap lies between green light to near IR light as shown in the middle of Fig. 2.6. II-VI materials' application can be found in Mercury Cadmium Telluride (MCT) detectors in focal plane array to image thermal radiation from an object [71].

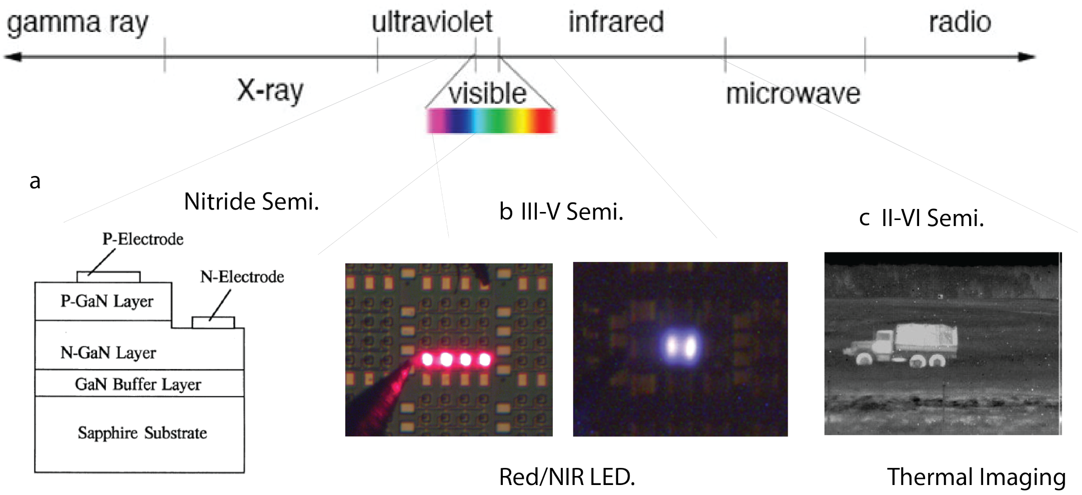


Figure 2.6: (a) First blue LED using wide band gap compound semiconductor. (b) Red and NIR (900 nm) LED made of III-V materials (in-house made by Seokhyeong Lee). (c) Thermal image taken by II-VI compound semiconductor Focal Plane Array. Reproduced with permission from [70, 71]

2.3.2 History of III-V Semiconductor

A brief history of III-V optical devices starts with GaAs and Optoelectronics in 1960s; The development of Gallium Arsenide (GaAs) as a semiconductor material in the 1960s marked a significant milestone. GaAs proved to be excellent for optoelectronic devices due to its direct bandgap and high electron mobility. Early applications included light-emitting diodes (LEDs) and semiconductor lasers [72, 73, 74]. In 1970s-1980s, III-V semiconductors, particularly Indium Phosphide (InP), played a crucial role in the development of the fiber optics revolution. InP-based materials were used to create high-performance optoelectronic components such as lasers and photodetectors, enabling the expansion of long-distance telecommunications via optical fibers [75, 76, 77]. Later on, the exploration of wide-bandgap III-V materials like Gallium Nitride (GaN) gained momentum. GaN, with its applications in blue and ultraviolet LEDs, later became a key technology for solid-state lighting and high-efficiency light-emitting devices [78, 79, 80]. Following years of late 1990s, advances in nanotechnology led to the exploration of quantum dots in III-V semiconductors. Quantum dots, with their unique quantum confinement properties, found applications in areas such as quantum dot lasers, single-photon emitters, and quantum information processing [81, 82, 83]. III-V compound semiconductors have become integral to the development of optoelectronic integrated circuits (OEICs). These circuits integrate optical and electronic components on the same chip, enabling high-speed communication and signal processing in a compact form. III-V semiconductors continue to be at the forefront of emerging technologies. They are critical for the development of next-generation optoelectronic devices, including those for quantum information processing, terahertz technology, and advanced imaging.

2.3.3 Future of III-V Semiconductor

The future of III-V compound semiconductors holds significant promise, especially concerning their integration into various technological realms. In electronics, III-V materials are seen as potential candidates to extend beyond the limitations of silicon in advanced electronics. Their high carrier mobility and superior electronic properties could enable faster and more efficient transistors, potentially overcoming some of the scaling challenges posed by silicon-based technology. In optoelectronics field, there’s a growing interest in hybrid integration, combining III-V materials with silicon. By integrating III-V materials onto silicon substrates, researchers aim to harness the advantages of both materials, enabling the creation of high-performance devices while leveraging the existing silicon manufacturing infrastructure. III-V semiconductors play a crucial role in enabling high-frequency and high-speed communication technologies. As the demand for faster data transmission increases with technologies like 5G and beyond, III-V compounds are likely to find expanded use in photonics and high-speed communication devices. Also, III-V materials are central to the development of quantum technologies, including quantum computing, sensing, and communication. Their unique properties at the quantum level make them promising candidates for building qubits and other essential components in quantum information processing.

On the other hand, advancements in manufacturing techniques are essential to scale up the production of high-quality III-V materials at a cost-effective level. Techniques such as molecular beam epitaxy and metalorganic vapor phase epitaxy are being refined to achieve large-scale production [84, 85]. Integrating III-V materials with other semiconductor platforms requires overcoming challenges related to material compatibility, thermal expansion mismatches, and interface engineering. Ongoing research focuses on developing techniques for seamless integration while preserving the unique properties of III-V compounds. Moreover, cost remains a significant factor inhibiting widespread adoption. Efforts are ongoing to develop cost-effective manufacturing processes and explore alternative substrates to make III-V technologies more economically viable.

Apart from traditional applications in electronics and optoelectronics, III-V compounds are finding use in diverse fields such as sensing technologies, biomedical devices, and even in emerging areas like flexible electronics due to their unique properties and performance advantages. Overall, the future of III-V compound semiconductors is promising, driven by ongoing research efforts aimed at enhancing their integration, improving manufacturing techniques, and exploring novel applications across a wide spectrum of technological domains.

Chapter 3

Waveguide integrated Black Phosphorus mid-IR modulator and photodetector

[1] Black Phosphorus (BP) exhibits distinctive optical characteristics that have found extensive application within various optoelectronic domains. Leveraging its pronounced tunability in optical absorption and layer-dependent band structures, BP facilitates highly efficient and wide-ranging optical photodetection and modulation spanning the spectrum from visible to mid-infrared (mid-IR) wavelengths. Despite rapid advancements in integrated mid-IR technology, driven by escalating requirements in optical communication and molecular sensing, prevailing mid-IR devices predominantly rely on bulk materials, presenting compatibility challenges with Si-photonics platforms. In this study, we present a novel integrated mid-IR modulator employing black Phosphorus, seamlessly integrated with a Si-waveguide. The modulator operates under the influence of the gate electric field, enabling the achievement of substantial optical modulation depths. Experimental results demonstrate an optical modulation depth of 8 dB/mm, while simulation models project a modulation depth as high as 61 dB/mm. This integration signifies a crucial advancement towards bridging the gap between mid-IR functionalities and Si-photonics, potentially unlocking a new paradigm for mid-IR device engineering and application in integrated optical systems.

---

[1] Contents in this chapter include modified figures and descriptions from [86], with permission. Copyrights 2022 Ruoming Peng, University of Washington.

3.1 Introduction

Mid-Infrared (mid-IR) technology has drawn significant attention recently. Due to the increasing demands for the wide branch of applications, including free-space communication, bio-chemical detection, gas sensing, and astronomic distance galaxy detection/studies [87, 88, 89], varied novel photonic devices have been proposed where high-photodetection ability and modulation depth and speeds are highly desirable. For the free-space communication, the atmosphere offers several transparent windows in the mid-IR regime, mainly at 3-5 μm and 8-14 μm. Moreover, the development of mid-IR photonics will also benefit the research in chemistry and biology since various chemical molecules and gases have strong vibration and rotation absorption in the mid-IR regime, so-called molecular finger print regime. Although the mid-IR photonics poses crucial importances with useful applications, the conventional mid-IR devices still rely on the brittle compound semiconductors including InAsSb and HgCdTe which are costly and difficult to be integrated with Si-photonic platform. As an alternative, the 2D materials such as bP and graphene have demonstrated their seamless integration and compatible fabrication processes for photo detection and electro-optical modulation in the mid-IR regime. Especially, bP has advantages of its moderate band gap and its thickness dependence, giving direct bandgap ranging from 0.3 eV in the bulk to 1.5-2.0 eV in monolayer (the corresponding optical wavelength range is 0.62-4.3 μm). When alloyed with arsenide (black-AsP), it can further shrink the bandgap and extend optical wavelength range up to 8.5 μm with high concentration doping of arsenide [90, 91], covering wide transparent windows in mid-IR regime. Moreover, bP's high mobility for both carriers even at room temperature (1000 cm^2/(Vs)), suggests its candidacy in high speed electronic and optical devices.

3.2 Mechanisms for Electro-optical Modulation

3.2.1 Quantum Confined Franz-Keldysh Effect

Semiconductors under the influence of the electric field exhibit the changes in the band transitions, also known as Franz-Keldysh (FK) effect [92]. This effect is first observed in the CdS crystal where the optical absorption around the band edger was extended to longer wavelength, indication of the decrease of the bandgap. The physics behind this phenomenon is the finite overlap between the electron and hole wavefunctions each of which extends into the gap region due to the large applied electric field. The applications of FK in bulk semiconductors can be found in the electro-absorptive modulators using SiGe, GeSn, and InGaAs integrated on Si platforms [93, 94, 95]. Atomically thin 2D materials show more significant field modulation with added vertical confinements. Black Phosphorus, being a 2D material, will form the vertical quantum well and induce the discrete energy levels denoted as Ec1, Ec2 ... and Ev1, Ev2 ...[**?**]. From the theoretical calculation, the bandgap of the bP will decrease as we increase the electric field as it has been demonstrated that a large built-in field induced by layers of potassium atoms can even close the bandgap and form an anisotropic semimetal state of black Phosphorus [96]. According to the first principle calculation, the conduction band and valance band will touch at the field strength of 4.5 V/nm. With even larger field applied, bP can have topological phase where the band inversion will occur. Although those extreme cases of band deformation are not used, the quantum confined Franz Keldysh effect (QCFK) is still interesting and useful phenomenon for optical absorption modulation. In addition to the red shift of the absoprtion edge with induced electric field (FK effect), the discrete energy level due to the vertical confinement will also contribute to the effect of the field modulation. Considering the flat band condition, the split energy level such as Ec1, Ec2... and Ev1, Ev2 ... can only be accessed with the same band indices due to the symmetry of the wavefunctions. For example, transition between Ec1 and Ev2 is forbidden becasue the wavefunctions of them are orthogonal with each other resulting in total transition probability of $\tilde{0}$, whereas the transition between the same indices, such as Ec1 and Ev1 have finite wavefunction overlap, and corresponding transition moments. However, the applied electric field can perturb the wavefunction in the perpendicular dimension (z in Fig. 3.1). As a result, the optical absorption in few-layer bP will also change becoming the key principle for our electro-optic modulation. Considering p-doped nature of bP, the material has initial band bending due to the built-in field of the flake as shown in Fig. 3.1.

3.2.2 Burstein-Moss Effect

The Burstein-Moss effect is a change in bandgap of the semiconductor resulting from band filling, according to Pauli's exclusion principle. By applying an external gate voltage, the increase of charge carrier density reduces the available unoccupied electronic states, leading to decreased optical transition probability. Figure 3.2 shows the band diagram explaining the prohibited band transition with elevated Fermi level [97].

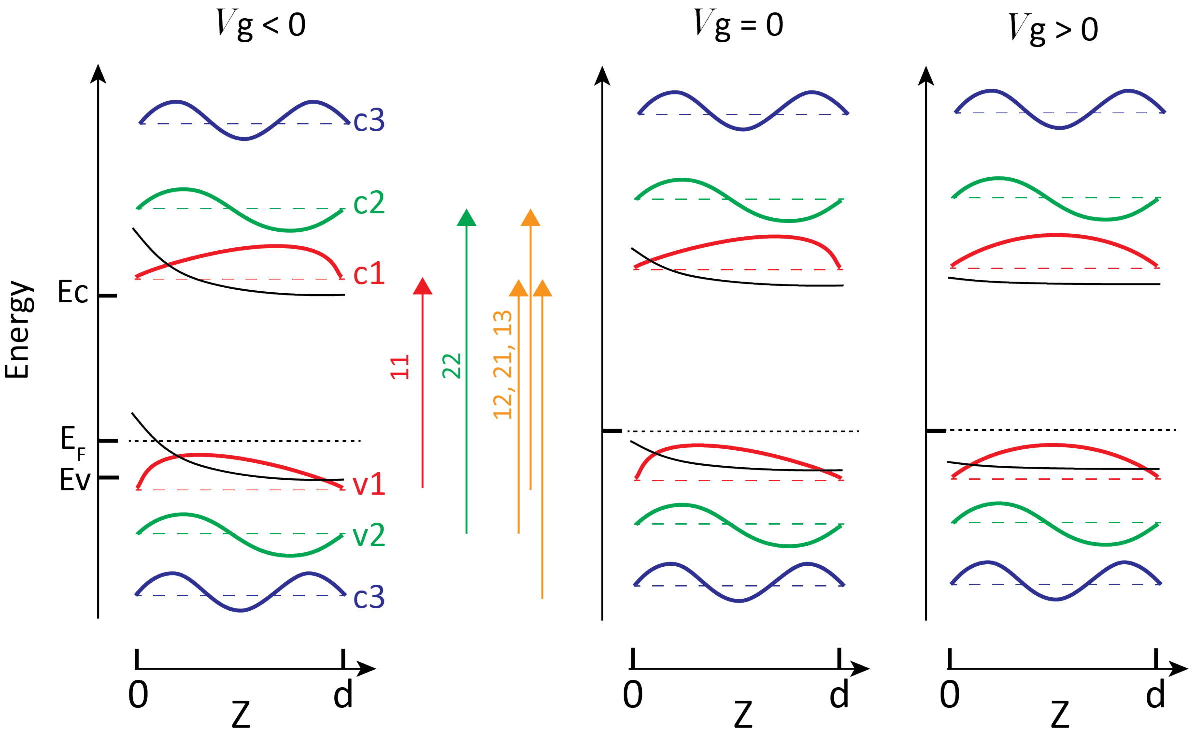


Figure 3.1: Schematics of enegy band diagram of bP under three bias conditions: Vg < 0, Vg = 0, Vg > 0. The colored-dash lines mark the energies of each sub-band and the solid lines depict the wave function in each sub-band. The black dotted line marks the Fermi level. The red/gree/yellow vertical arrows show possible transitions that can contribute to the modulation level extrema.

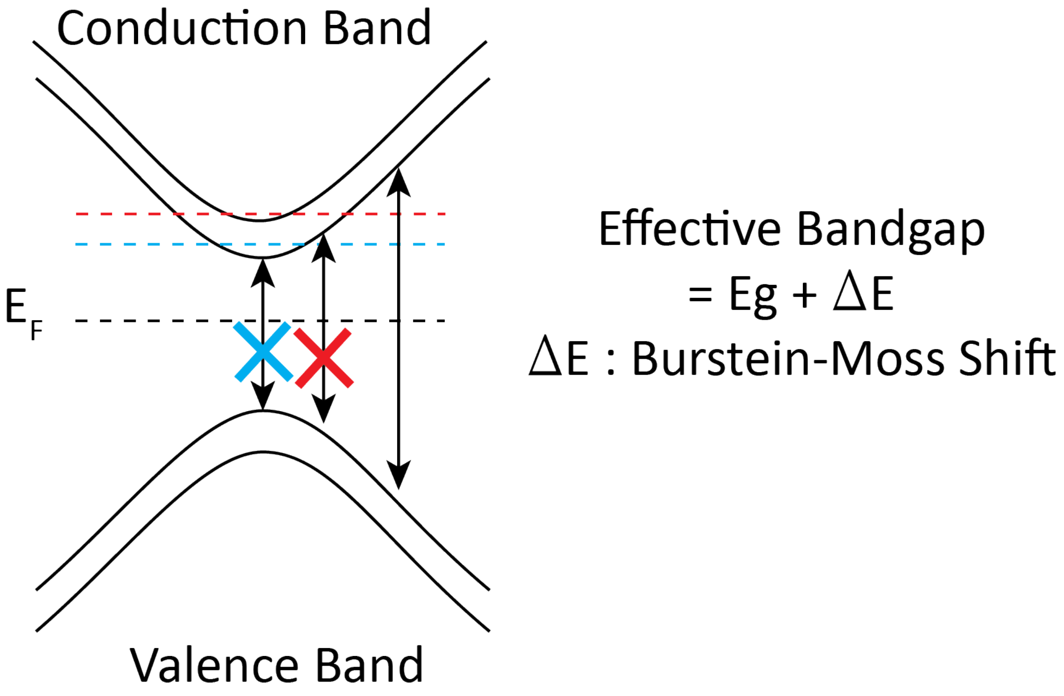


Figure 3.2: The bP-PPT array is capable of multispectral infrared imaging and is programmable for in-sensor computing. The array can be programmed remotely using optical control signals and locally using electrical gate voltages.

3.3 Numerical Simulation for Integrated Black Phosphorus Modulator

The complex refractive index of black Phosphorus (bP) can be extracted from the free-space transmission measurement results [[66]]. The bP devices can be considered to be stacked multilayers shown in Fig. 3.3(a). By importing the refractive index of each layer, we can construct the transfer matrix for the material interface (M) and the optical propagation (P).

$$M = \frac{1}{t_{21}} \begin{bmatrix} t_{12}t_{21} - r_{12}r_{21} & r_{21} \\ -r_{12} & 1 \end{bmatrix} \tag{3.1}$$

$$P = \begin{bmatrix} e^{j\phi} & 0 \\ 0 & e^{-j\phi} \end{bmatrix} \tag{3.2}$$

By applying the transmission spectrum of the devices shown in Fig. 3.3(b), the extinction coefficient of the bP thin film at different gate voltages is extracted. The shaded area here in the figure corresponds to the transparent windows in the mid-IR regime. The extinction coefficient of bP in this range is strongly modulated by the electric field becoming useful for free-space communication applications.

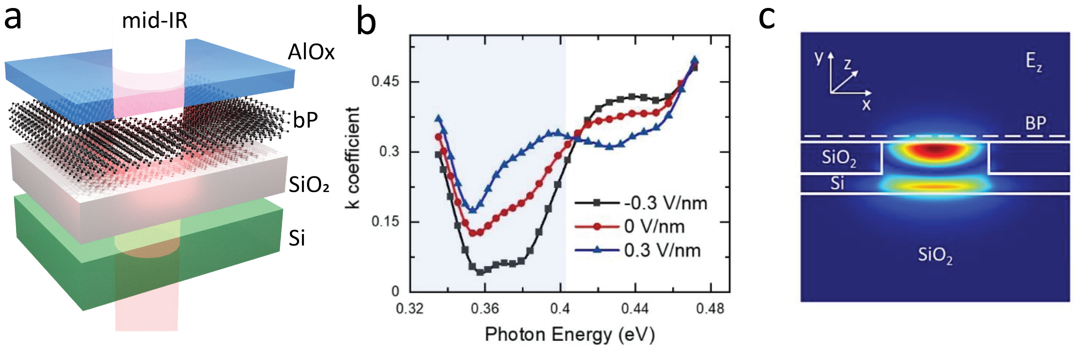


Figure 3.3: Simulation of the integrated bP devices. (a) Schematic structure of the free-space devices. Considering the multilayer structures, the extinction coefficient can be calculated from the transmission spectrum. (b) The calculated extinction coefficient of the bP thin film at different gate voltages. The optical absorption of bP is strongly modulated by the electric field especially in the optical spectrum range around 0.38 eV. (c) The simulation of the waveguide integrated device. The anisotropy of bP has been included in the simulation that the TM mode of the waveguide has a large mode overlap with bP thin film.

Considering this strong modulation of bP properties, we simulate the integrated bP-waveguide devices. We design the Ridge waveguide structure; the width is 1.6 μm with a total height of 600 nm. Considering the anisotropic property of BP, the armchair direction of BP is aligned with the Ez component of the TM mode in the waveguide, which we expect most interaction with the mode of the waveguide, and more modulation could be achieved. We construct the anisotropic matrix of BP and simulate the absorption of the fundamental TM mode as indicated in figure 3.4 (c). With the electric field ranging from -0.3 V/nm to 0.3 V/nm, the modulation depth can reach 40 dB/mm. the integrated bP devices show improved performance compared to other types of materials with integrated photonic platforms including graphene, Ge-Si, GeSn/SiGeSn MQW, Si free carrier absorption[**?**, 98, 99, 100, 101]. Also, by adopting the sandwiched photonic design for the graphene integrated devices, the modulation depth can be further enhanced. Among varied types of 2D materials, bP integrated devices have the strongest modulation with very low loss in the mid-IR region. According to the simulation of the waveguide integrated bP devices, 5 dB modulation with less than 1 dB insertion loss can be realized with 90 μm bP which meets the requirements for commercial applications.

3.4 Measurement Schematics

To measure the optical modulation of the integrated bP devices, the butt coupling setup is then introduced which can support the broadband transmission with high coupling efficiency. We used a mid-IR optical parametric oscillator (Firefly-IR, M Squared) as the light source outputting in the spectral range of 2.5-3.7 μm. As shown in figure 3.5, the tunable mid-IR light is coupled to mid-IR fiber which is connected to the end of the waveguide devices. To increase the coupling efficiency, the inverse tapered structures at the end of the waveguide are used to overcome the mode mismatching between the fiber and the Si waveguide. With the help of this tapered structure, the mid-IR light can be efficiently coupled from fiber to the waveguide and propagates through the bP devices. Then the light is collected by another mid-IR fiber connected with a photodetector for the transmission measurement. According to the SEM image of the device, the end of the taper is around 150 nm which will have the divergent mode profiles close to our mid-IR fiber. For the particular device, the optical transmission is calibrated which has transmitted signal several magnitude orders larger than the noise

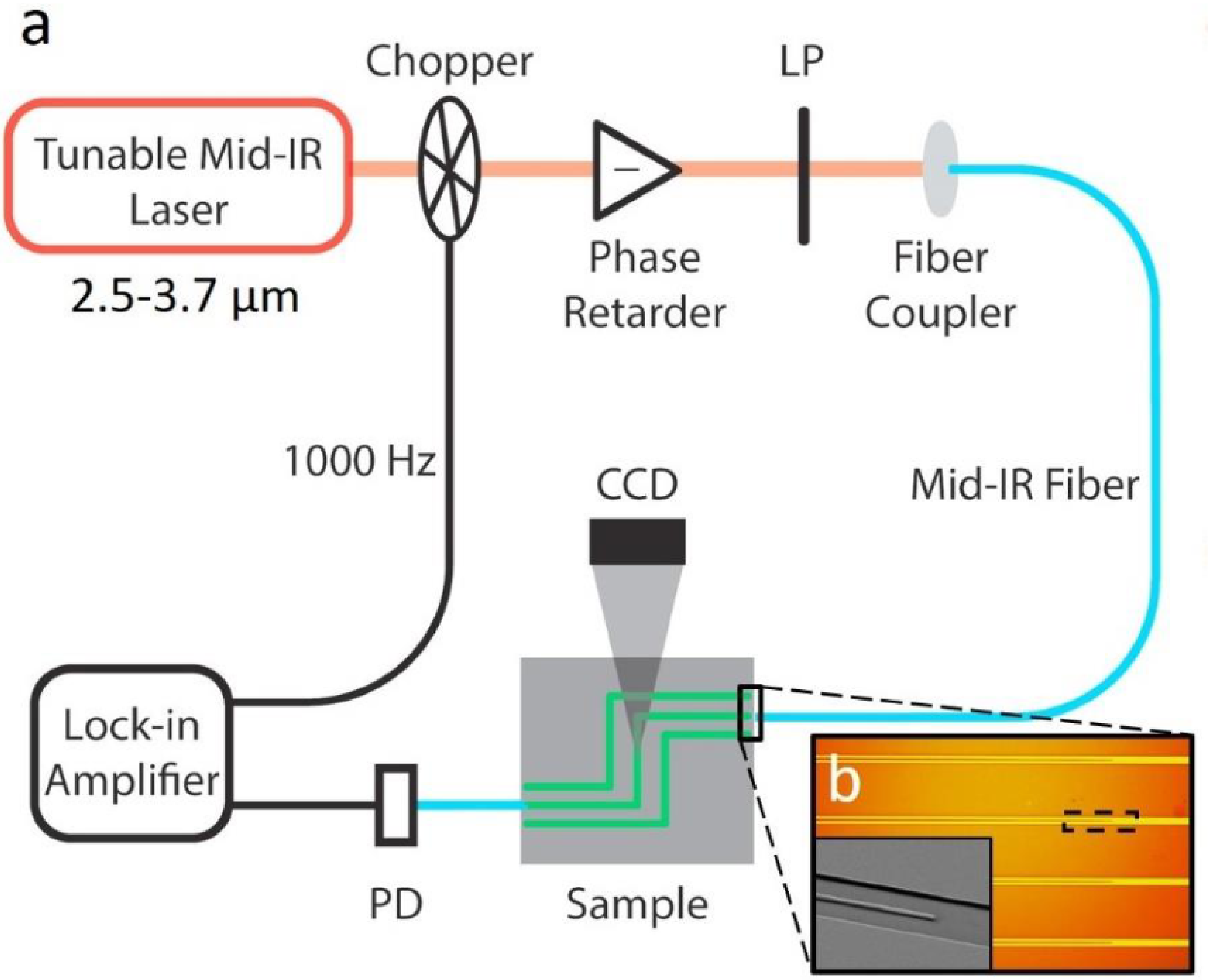


Figure 3.4: Measurement Schematics (a) The setup design for the photodetection and modulation measurements of integrated bP device. The phase retarder with a polarizer is used to control the output polarization of fiber to be aligned with the TM mode of the Si waveguide. (b) The optical image of the inverse tapered structures. Inset: SEM image shows the zoom-in feature of the tapered end where the width of the tapered end is about 150 nm

3.5 Integrated Black Phosphorus mid-IR Detector

To characterize the device performance of bP devices, we first apply the bias voltage to measure the photocurrent when incident the mid-IR light. By scanning the optical wavelength, we measure the photocurrent/responsivity spectrum of the bP devices at bias voltage 0.2 V. From the responsivity spectrum of the bP device, there are three resonances which are corresponding to the band transitions of the bP thin film. Considering the 14-nm bP thin film, there are four possible band transitions located within our light source range. According to the band transition diagram shown in fig. 3.5 (a), the first and third responsivity peaks are related to E11 and E22 transitions, while, the hybrid transitions of E12 and E21 cause the small bump in between. These hybrid transitions should be symmetry forbidden. However, the initial doping and band bending inside bP break the symmetry and enable the hybrid transitions. Here, we demonstrate an integrated bP photodetector for mid-IR applications with external responsivity up to 2.5 mA/W. At a certain wavelength (λ=3 μm), the power-dependent optical response of integrated bP devices is characterized. The photocurrent of bP devices shows the linear dependence with the incident power which is crucial for photodetection.

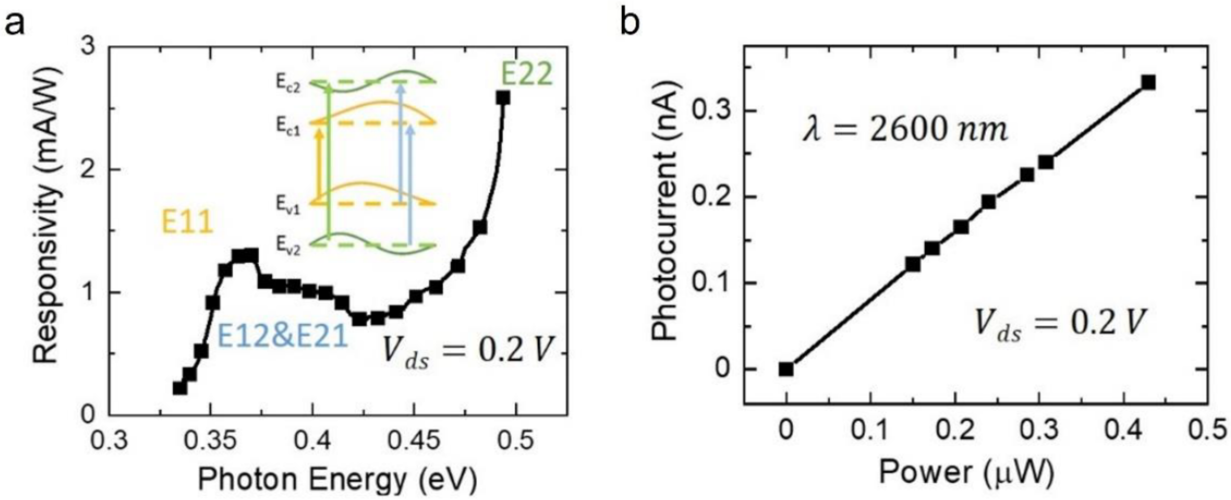


Figure 3.5: Integrated bP device for photodetection (a) The responsivity spectrum of bP photoresponse. There are three resonances in the responsivity spectrum which are corresponded to the subband transitions of bP thin film. (b) The power-dependent photocurrent measurement of integrated bP device. The device can achieve the photodetection at sub μW regime.

3.6 Integrated Black Phosphorus electro-optical Modulator

We demonstrated the broadband mid-IR modulator by integrating the bP with Si waveguide and characterized the modulation depth of the integrated device. For the gate modulation, we connect the gate electrode with the Si rib waveguide. As the gate voltage is applied, the electric field will be formed through the Si to modulate the bP thin film. Considering the high-k dielectric layer (AlOx), the optical absorption of bP is efficiently modulated, and consequently, the transmission of the integrated bP device is modulated by the external electric field. The 2D plot in fig. 3.6 (a) shows systematic measurement results of the modulation level when both wavelength λ and gate voltage Vg are continuously scanned. The modulation M always refers to the transmission when the gate voltage is 0 V: $M = [T(V_g) - T(0)]/T(0) = \Delta T/T(0)$. According to the modulation results, the transmission of the devices is depending on the applied gate voltage. At a certain wavelength, the modulation depth is up to 8 dB/mm. For characterization, we extracted the modulation of the transmitted signals at two different gate voltages (±10 V). As shown in fig. 3.6 (c), the bP device shows a broadband modulation from 2800 nm to 3400 nm. The reason that measured modulation is weaker than the simulation, can be explained by the fact that the gate voltage induced through the intrinsic Si back gate cannot efficiently modulate the bP. With more efficient gate control such as ionic liquid, the modulation depth can be further enhanced. Also, the waveguide region of the device is not a perfectly flat substrate that

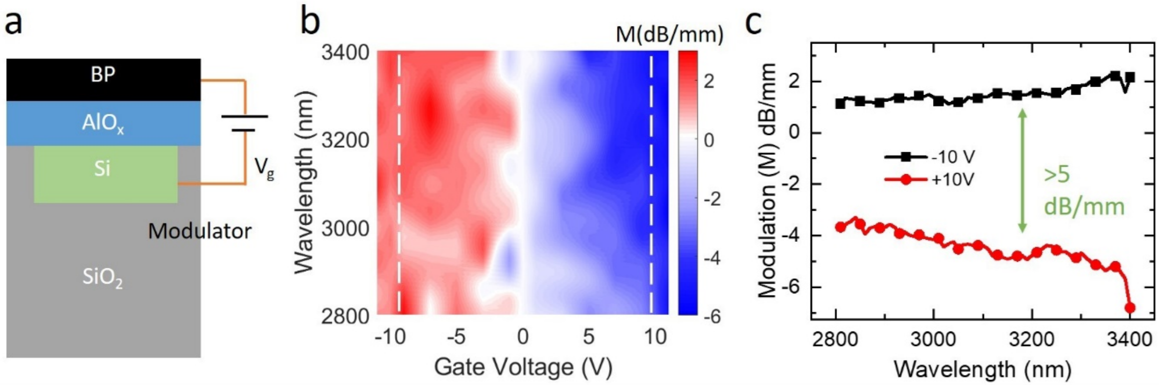


Figure 3.6: Integrated bP device for optical modulation. (a) The measurement schematic for the integrated bP device. This device can be operated as both the photodetector and modulator. (b) The modulation level is measured as the function of the wavelength and applied gate voltage, featuring the max modulation of about 8 dB/mm. (d) Measured modulation at gate voltage ±10 V, the device shows broadband modulation in the mid-IR range from 2800 nm to 3400 nm

EO Material	Max MD (dB/mm)	Wavelength (μm)
Si (free-carrier)	0.5	2
Graphene (Sim.)	8.2	2.05-2.45
Graphene (Exp.)	7.8	2.05-2.45
BP (Sim.)	> 40	3-3.4
BP (Exp.)	8	2.8-3.4

Table 3.1: Overview of the mid-IR Electro-absorptive modulators

induces the inhomogeneous broadening of transition resonances and significantly degrades the performance of the bP modulator. The more careful planarization method such as the Chemical Mechanical Polishing (CMP) and the device geometry with the top gate electrodes can contribute to the stronger modulation results. Although the modulation is relatively weak compared to the simulation results, the modulation depth measured in our integrated bP device is already comparable to the graphene mid-IR modulator (sandwiched in the ChG waveguide)[102]. As shown in Table 3.1, the modulator device based on the free-carrier absorption of Si only has a modulation depth of about 0.5 dB/mm which is much smaller than the results measured in the integrated bP device.

To compare with the other integrated modulator devices, our bP devices already show improvement in the modulation depth. The current performance is still below the early simulation results or other theory predictions, 40 dB here, there is much freedom we could improve in this system by optimizing the fabrication process and protecting the bP from degradation. The integrated bP devices show advantages over other types of optical modulators that the footprint of the modulator can be further reduced with a similar modulation depth.

3.7 The Integrated platforms for high Modulation Depth

Considering the waveguide integrated devices, the modulation is far below the simulation results. There are a few problems with these waveguide devices. First, only the evanescent field is leaked into BP, the light-matter interaction is very small. As result, it requires a 180 μm long device to achieve enough modulation. Secondly, the etched waveguide region must be planarized for the following transfer process of BP. Typically, the SiO2 is deposited to fill the etching region and planarize the device. However, the thermal evaporated SiO2 is very lossy in the mid-IR range and causes additional internal loss in the waveguide devices. The last but most important thing is that the planarized region cannot be perfectly smooth, the transferred BP on this region will be very ununiform. This will degrade the optical property of the BP because 2D materials are very sensitive to the surface interactions and the local strain.

3.7.1 Multi-pass Waveguide Design

Considering the exfoliation of the bP flake, it would be challenging to achieve >100 μm size device. With the flake size of 50 μm, the interaction length would be limited which results in a small modulation depth. Fortunately, by taking the advantage of Silicon photonics, enough modulation depth can be achieved within a 50 μm × 20 μm footprint. By designing the multi-pass waveguide as shown in figure 3.7, the light would propagate through the device with 5 passes. According to our simulation, 5dB modulation depth with less than 1 dB insertion loss can be achieved in this multi-pass waveguide device.

3.7.2 Ionic Liquid for stronger Field modulation

As discussed above, the waveguide device suffers the weak back gate modulation due to the intrinsic silicon. Also, the atomic layer deposition of AlOx for the gate and material protection will add additional insertional loss to the waveguide device. With the complicated fabrication process, the bP thin film will expose more to the ambient condition which can degrade the optoelectronic properties of bP. To achieve the stronger gate modulation, the ionic liquid(iL) BMIM-BF4 is deliberately chosen for ambient and room temperature measurement because it is reported that the air degradation of BP thin film can be protected for tens of days in the air by directly applying the iL on the BP flake[103]. The applicable voltage window of the iL is from -1 V to 1 V which can modulate the carrier concentration of the thin film in the order of $10^{1}3/cm^{2}$. It can achieve more than 10 times of the modulation range by using the AlOx dielectric layer.

The electrical characterization of the test iL device shows a strong gate modulation of bP conductivity with a bias voltage of 20 mV (Figure 3.7). The IV characteristics indicate the effective doping of bP by ionic liquid with an on-off ratio over 1000 and the charge neutral point at 0.3 V. The modulation of optical transmission for the iL device was demonstrated with more than 10 dB/mm and a wide spectral range from 2950 nm to 3150 nm.

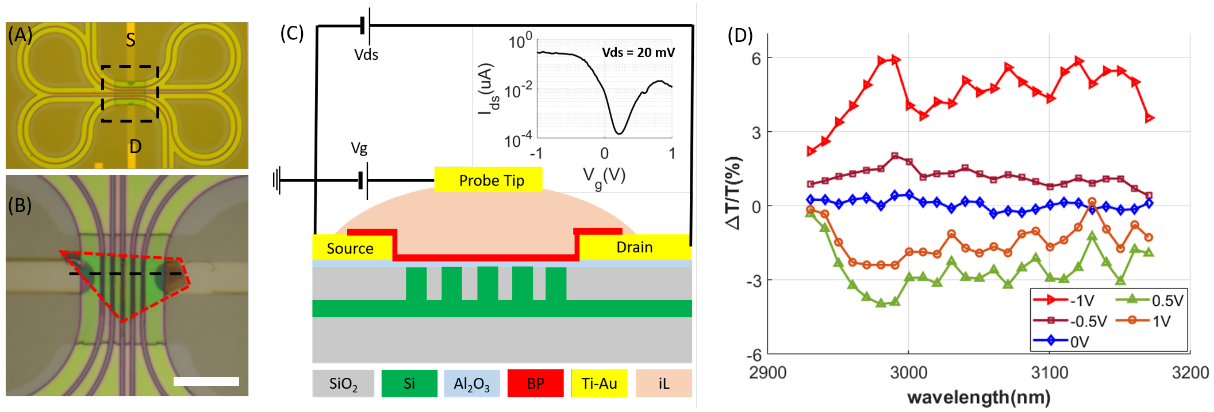


Figure 3.7: Characterization of the iL-gate devices. The optical images of (a) the multi-pass waveguide and (b) the bP-integrated waveguide. BP is indicated with the red-dashed line. The green square region is planarized with SiO2. Scale bar: 20 μm. (c) The cross-sectional schematics of the iL-gate device. Inset: Field-effect modulation of BP conductance. (d) Transmission changes normalized to that of zero gate voltage scanned around 3-um at different gate voltages. The maximum modulation occurs as 10% at 2.98 μm between gate voltages of 1 and 0.5 V.

3.7.3 Plasmonic waveguide for enhanced light-matter interaction

To further enhance the performance of the optical modulation, the plasmonic nanogap geometry is also introduced. A more compact footprint and higher modulation depth can be achieved by using a metal-insulator-metal plasmonic waveguide (MIM-WG) structure because of its strong field enhancement, which increases light-matter interaction[104]. The plasmonic guided mode for 3.3-μm light is simulated with FEM with and without BP integrated on top of the waveguide, according to the geometry shown in Figure 3.8 (a) with the varying gold thickness (t_{au}) and slot width (w_{slot}). The refractive index (n) and anisotropic extinction coefficient (k) of BP are cited from reference[104]. After the plasmonic mode is simulated, the absorption coefficient (dB/μm) is calculated from the imaginary part of the complex effective mode index.

The modulation strength is calculated from the change of the absorption coefficient when assuming the extinction coefficient of BP is modulated between 0.3 and 0.05. The results show that only 5 μm long MIM-WG with 20 nm slot width and 20 nm gold thickness is sufficient to achieve 5 dB modulation depth (M). However, as the gap size and thickness of gold reduces, the insertion loss also increases. To optimize for a high modulation depth (M) while minimizing the insertion loss (IL), we define a figure of merit (FOM) as: FOM=M/IL with the restriction that M>5 dB and device length less than 10 μm. With these conditions, we find 50-nm thick gold is optimal for achieving M=5 dB, IL=2.5 dB with a BP flake of 8 μm long.

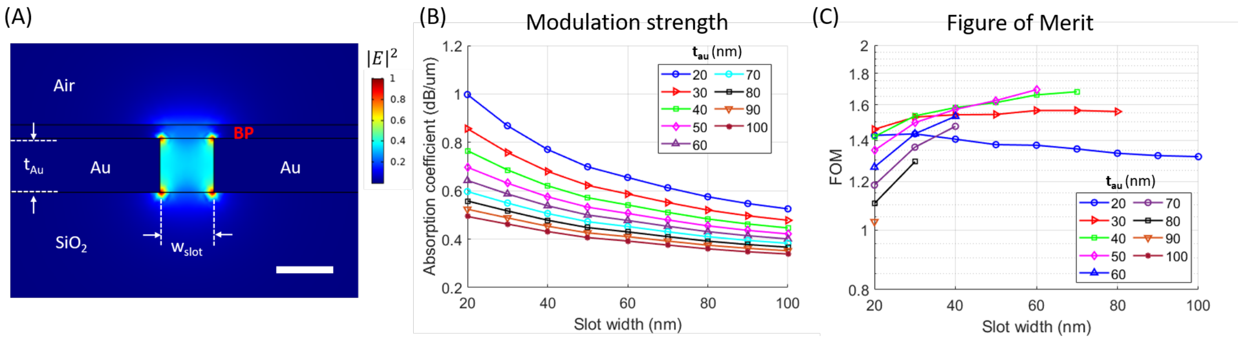


Figure 3.8: Integrated plasmonic devices for enhanced optical modulation (a) Schematics of MIM-WG simulation. Scale bar: 40 nm. The electric field is highly confined near the edges of gold and BP nearby, enhancing light-matter interaction. (b) The modulation strength, defined in the main text, is calculated with varying gold thickness and slot width. (c) Figure of Merit, defined in the main text, considering both modulation depth and insertion loss

3.8 Conclusion and Outlook

To summarize, we have demonstrated the waveguide integrated bP photodetector and modulator for the mid-IR applications. For the photodetection, a broadband photoresponse is demonstrated with the responsivity up to 2.5 mA/W. Meanwhile, we achieve the optical modulation with the modulation depth of 8 dB/mm in the experiment and 40 dB/mm based on the simulation. We believe that the large modulation predicted by our simulation is promising as we introduce the top gate geometry to achieve more efficient grate control and also planarize the waveguide region to reduce the inhomogeneous broadening of bP optical resonances. By taking advantage of the novel photonic design such as plasmonic waveguide, and photonic crystal structures, the modulation depth can be further enhanced which can easily meet the criteria for the commercial applications. The above results for the simulation and experiment of the integrated bP device indicate that bP has great potential in mid-IR photonics, especially for broadband photodetection and optical modulation.

Chapter 4

In-sensor computing using Programmable Black Phosphorus Phototransistor

[1] Electrostatic gating to modulate the electrical and/or optical properties of semiconductor is well-documented knowledge that is widely used in academic and industrial scenes. 2D materials, thanks to their drastic modulation of the electrical and/or optical properties compared to traditional 3D semiconductor materials, are considered promising materials for faster and smaller electronics and opto-electronic applications. Black phosphorus, among 2D materials, has unique advantages that it has direct bandgap tunable from 0.3 eV to 2.0 eV depending on its thickness, high electron/hole mobility, and broadband optical bandwidth. Leveraging its broadband infrared responses, arrays of bP photodetectors can be utilized for multispectral imaging, which acquires spatial images with spectral information. Multispectral imaging combined with artificial neural networks (ANN) has become a powerful tool for biomedical imaging [38], fresh food classification [65], and surface damage detection on industrial sites [106].

---

[1] The contents in this chapter are adapted with permission from: Lee, S., Peng, R., Wu, C. & Li, M. Programmable black phosphorus image sensors for broadband optoelectronic edge computing. Nature Communications 13, 1485 (2022)[105]. Copyright 2022 The authors. Nature Publishing Group

4.1 Introduction

4.1.1 Remote Programmable Infrared Vision Processor

With the explosive growth of data generated from miniaturized sensors, cameras, texts, and information in daily life, the importance of data processing at close vicinity of each sensor is drawing attentions[107, 108]. To overcome the data bottleneck from the serial processing of Von Neumann architecture, in-memory computing for parallel processing that combines memory elements and computing units at the same place is proposed[**?**, 109]. The advantages of the in-memory computing are large reduction of data traffic of redundant data, and relaxed communication delays between edge devices and a remote powerful computing center.

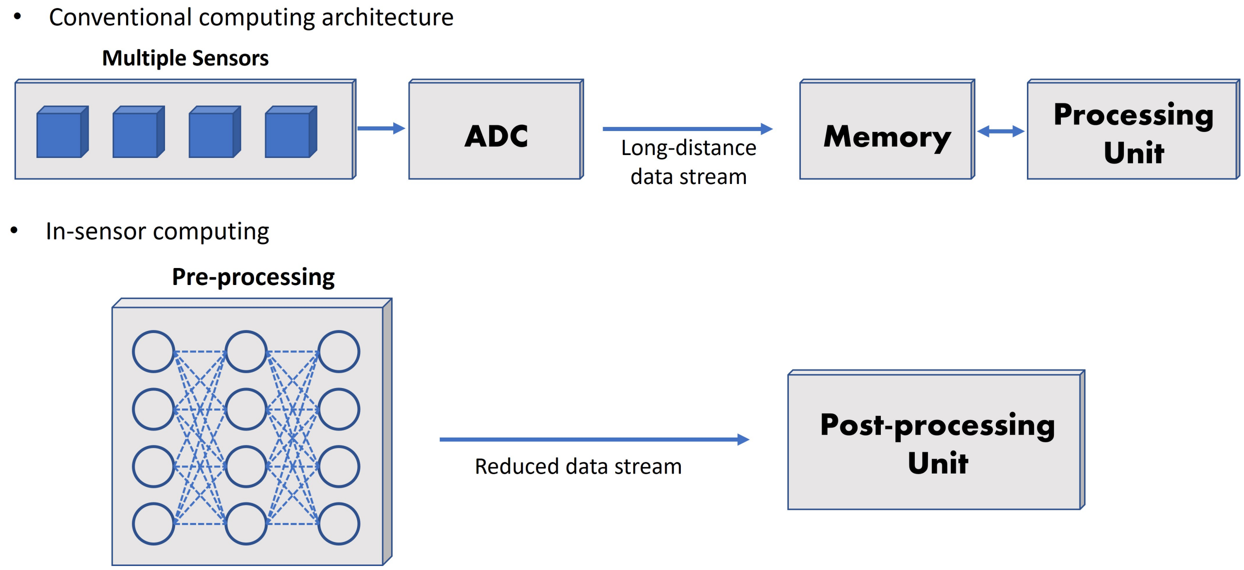


Figure 4.1: The difference between conventional Von Neumann computing architecture and in-sensor computing scheme. Pre-processing data at edge device significantly reduces the data loads as well as latency.

The preprocessing conducted at the edge devices is analogous to the human vision system, encompassing not only eyes and photoreceptors but also optic nerves and the brain to form an efficient vision processing system, as illustrated in Fig. 4.1. The emerging concept of in-sensor computing, involving preprocessing at the sensor end, has garnered attention through numerous research endeavors and demonstrations [110, 111]. This chapter demonstrates the in-sensor computing paradigm incorporating two fundamental components, namely an optical image detector and a subsequent processing unit, seamlessly integrated on a unified platform utilizing the black phosphorus phototransistor (bP-PPT). The programmable photodetector array exhibits the capacity to preprocess an image upon its acquisition by dynamically adjusting the responsivity of each pixel. By incorporating a parallel computing processor, this arrangement enables more sophisticated neural network computations. The optically programmable bP-PPT provides the added benefits of simultaneous parallel programming of individual pixels and remote programmability, as depicted in Fig. 4.2.

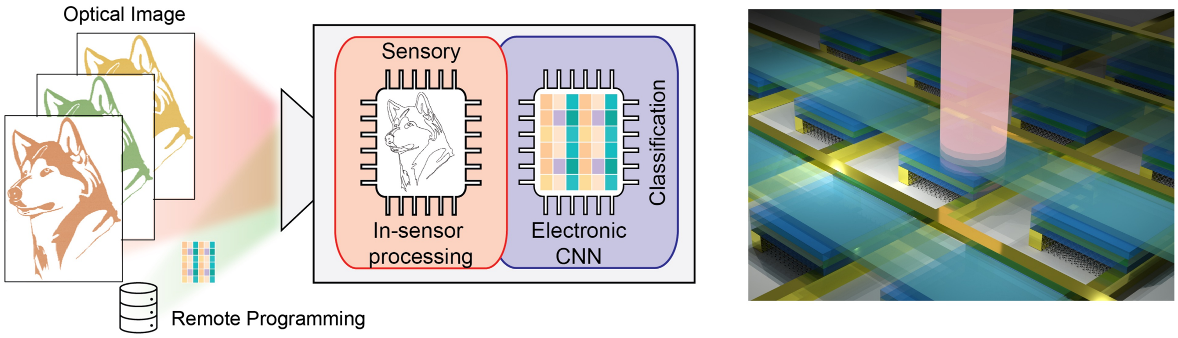


Figure 4.2: The bP-PPT array is capable of multispectral infrared imaging and is programmable for in-sensor computing. The array can be programmed remotely using optical control signals and locally using electrical gate voltages.

4.2 Device Structure and Working principle

The field effect transistors made out of black phosphorus (bP) thin films with various geometries from different groups have revealed the strong electrical/optical tunability with gate electric field. For electrical conductivity, black phosphorus has large on-off ratio of 10^5 with a high mobility for both electrons and holes with 1000 cm^2/Vs in room temperature and 10^5 cm^2/Vs in cryogenic temperature [63]. Multiple research groups have extensively explored and managed the optical properties of black phosphorus across a wide wavelength spectrum, spanning from visible to Mid-IR. The manipulation of both electrical and optical characteristics is achieved through gate geometry modulation. The bP-PPT device is engineered with a dielectric stack composed of $Al_2O_3/HfO_2/Al_2O_3$ (AHA), allowing programmability by storing charges in the hafnium oxide layer (Fig. 4.3 a). Indium-tin-oxide (ITO) serves as the transparent top gate electrode, providing optical access to the bP-PPT. The band alignment of multiple layers in the device is depicted in Figure 4.3 c. The layer structure and material selection are designed to enable tunneling of charges (electrons or holes) from the bP channel through the thin Al_2O_3 barrier layer, storing them in the HfO2 layer and effectively modulating the bP channel through field effect. The electron affinity difference between Al_2O_3 and HfO_2 ($\chi_{Al_2O_3}$=1.5eV, χ_{HfO_2}=2.5eV) creates a high tunneling barrier for the retention of trapped charges. The energy difference between the stored charges and the conduction band of HfO_2 determines the storing energy to be 1.25 eV, allowing optical control of the stored charges. Visible light ($\lambda < 0.992$ µm) can remove the charges, whereas infrared light in the telecommunication band or longer wavelengths cannot. An optical image of the 4 x 3 array of bP-PPT devices, fabricated from a single bP flake with 11 nm thickness, is shown in Figure 4.3 b. The collective measurement results of the source-drain current (I_{ds}) over gate voltage (V_G) sweeping are displayed in Figure 4.3 d. The hysteresis loop in the Ids–V_G curve, with a substantial memory window of 25 V in V_G, reflects the injection and storage of charges during the V_G sweep. The high charge storage density ensures effective control of the electrical conductance of the bP-PPT, achieving an on/off ratio > 200. Notably, the array exhibits excellent uniformity, with inter-device variation in the on/off ratio among nine devices being less than 8 % (inset, Figure 4.3 d).

4.2.1 Fowler-Nordheim Tunneling model

The working principle of the bP-PPT device can be modeled by the Fowler-Nordheim tunneling (FN tunneling), which explains the charge trapping and de-trapping mechanism by electrical and optical control, whose measurement result will be discussed in the following section. Here, rigorous analytic approach is taken to model the behavior of BP-PPT devices. Fig. 4.4 depicts the band alignment of the bP channel, Al_2O_3 tunnel layer, and HfO_2 charge storage layer. When a sufficiently large voltage is applied to the top ITO gate, the large electric field across the tunneling layer can lead to the FN tunneling. A tunneling current will be injected into the HfO_2 trapping layer, where charges are trapped at trapping sites with energy in the bandgap [112]. After the applied gate voltage is removed, these metastable trapped charges remain in the HfO_2 layer and induce effective gating to the bP channel, modulating its optical and electric properties. Fig. 4.4 b illustrates the band diagram under the built-in electric field by these trapped charges. Under optical illumination, the trapped electrons can be excited and escape from the trap site to tunnel back to the bP-channel,

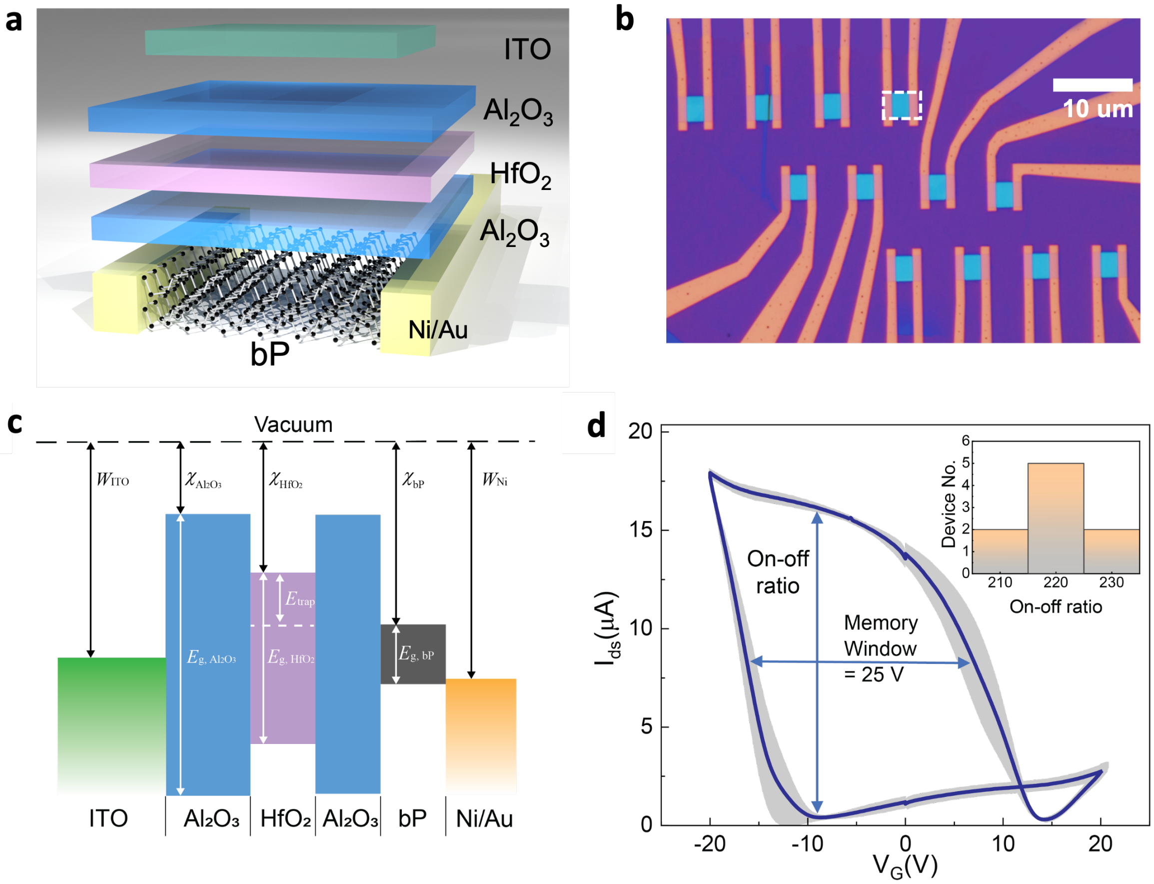


Figure 4.3: bP-PPT device working principle. (a) the graphical schematics of bP-PPT device. (b) The optical microscopic image of 4 x 3 array of bP-PPT devices. Light Blue square indicates a bP channel patterned in a rectangle. Scale bar: 10 um. (c) The band alignment of bP-PPT devices. (d) Collective gate swept IV curve of bP-PPT devices with large memory windows and on-off ratio. Inset: histogram of on-off ratios.

facilitated by the built-in field as in Fig. 4.4 b. The amount of tunneling charges can be precisely controlled by the optical power and pulse duration to realize 36 intermediate states. The electric field distributions with and without gate voltage are depicted in Fig. 4.4 c and d, for different charge densities in the charge storage layer (long-dashed line, short-dashed, and straight lines for the highest, moderate, and no charges, respectively). To estimate the electrical programming/erasing speed and optical programming speed, we calculate the FN tunneling current with a triangle barrier of bP/AHA/ITO device. The quantum tunneling transmission function TC(ξ) is first calculated using the Wentzel-Kramers-Brillouin (WKB) approximation, and is given by [62],[113],

$$TC(\xi) = \exp\left(\frac{-2}{\hbar}\int_0^{t_{ox}} \sqrt{2m_{ox}(V(x)-\xi)}\,dx\right) \tag{4.1}$$

where t_{ox} is the thickness of the tunneling layer (Al_2O_3), m_{ox} is the effective mass of the carriers in the tunneling layer, V(x) is the potential function of the triangle-barrier, and ξ is the energy of incident carriers referenced to the Fermi energy of the electrode. When the gate voltage is applied, the electric field across the tunneling Al_2O_3 layer is calculated using the capacitor model:

$$V_G = \frac{\sigma_1}{\epsilon_{Alo}}t_{blc} + \frac{\sigma_1}{\epsilon_{Hfo}}t'_{str} + \frac{\sigma_1+\sigma_{Hfo}}{\epsilon_{Hfo}}(t_{Hfo}-t'_{str}) + \frac{\sigma_1+\sigma_{Hfo}}{\epsilon_{Alo}}t_{tnl} \tag{4.2}$$

$$E_{ox} = \frac{V_{tnl}}{t_{tnl}} = \frac{\sigma_1+\sigma_{HfO}}{\epsilon_{ALo}} \tag{4.3}$$

where σ_1, σ_{HfO_2} are the charge density at the interface between the gate electrode and the blocking layer, and the charge density stored in the HfO_2 layer, ϵ_{AlO}, ϵ_{HfO} are the dielectric constant for Al_2O_3 and HfO_2, t_{tnl}, t_{HfO}, t_{blc} and t'_{str} are the thickness of tunneling layer, charge storage layer, blocking layer, and the position for the barycenter of the stored charge, respectively. We assume the charge density is uniformly spread in the HfO_2 layer so that the barycenter is assumed as the center of the HfO_2 layer. Accordingly, the tunneling current through the tunneling layer can be estimated as [113]:

$$J_{FN} = \frac{m_e q^3}{8\pi m_{ox} h q \Phi_B} E_{ox}^2 \exp\left(-\frac{4\sqrt{2m_{ox}}}{3\hbar q E_{ox}} q\Phi_B^(3/2)\right) \tag{4.4}$$

where q,h,m_e, and Φ_B are electron charge, Planck's constant, electron mass, potential energy barrier at the bP and Al_2O_3 interface, respectively. The stored charge density in the HfO_2 layer (σ_{HfO_2} (t)) (negative for electrons and positive for holes) is the integration of the tunneling current density over the operation time t,

$$\sigma_{HfO_2}(t) = \int_0^t J_{FN}(t')\, dt' \tag{4.5}$$

The stored charge density after the programming by gate pulse can induce the effective gating to the bP-channel. Meanwhile, those stored charges in the HfO$_2$ layer can form the built-in electric field across the tunneling layer and cause the charges leakage to the bP channel. The tunneling current of the leakage process depends on the tunneling coefficient and the carrier density regarding the process, which can be described by Tsu-Esaki formula [114]:

$$J_{leak} = -q\frac{dn_{stored}}{dt} = \frac{q}{4\pi^3\hbar}\int_{\xi_{min}}^{\xi_{max}} TC(\xi)g(\xi)f(\xi)\, d\xi \tag{4.6}$$

where $g(\xi)$ and $f(\xi)$ are the density of state of charge carriers and distribution function, respectively. For simplicity, we assume charge trapping in HfO$_2$ is dominated by one type of traps so simplify $g(\xi)$ with delta function, $g(\xi) = n_{stored}\delta(\xi - \xi_{trap})$. $f(\xi)$ is Fermi-Dirac distribution function. Then, since the tunneling coefficient depends on the built-in field and the stored charge density, we have:

$$\begin{aligned} -\frac{dn_{stored}}{dt} \propto & \{TC(\xi_{trap}, E_{built}(n_{stored})) * n_{stored} + TC(\xi_{th}, E_{built}(n_{stored})) * n_{th} \\ & + TC(\xi_{opt}, E_{built}(n_{stored}) * n_{opt}\} \end{aligned} \tag{4.7}$$

where the first and second terms of the right-hand side are related to the leakage process without any external illumination. The first is the direct tunneling due to built-in field from the traps of the stored charges and the second term is thermally excited charges from traps to conduction band with energy ξ_{th} and contributes to tunneling. In the third term, we also consider the optical excitation of the stored charges with the photon energy $\hbar\omega$ and the excited electron energy $\xi_{opt} = \hbar\omega + \xi_{trap}$, which are related to the tunneling coefficient TC, while the stored charge density in the storage layer also affects the built-in field that assists the tunneling process.

From equation 4.8 , the stored charges can screen the effective electric field in the tunneling layer. Hence, the FN tunneling current is suppressed with increasing density of stored charge, which we calculate

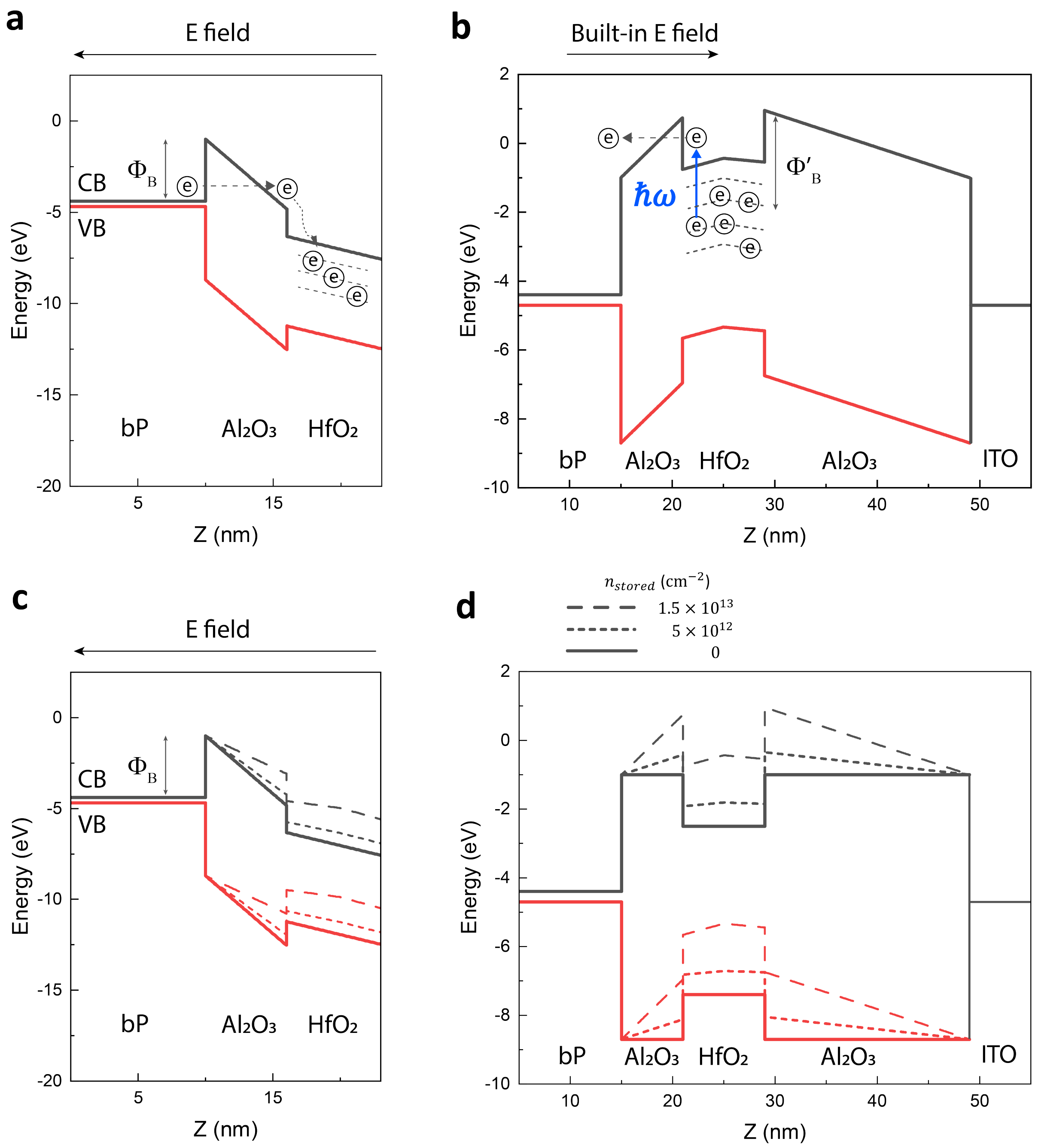


Figure 4.4: The working principle of programming and erasing of the bP-PPT. (a) Band diagram of bP/AlO/HfO layer when 18 V top gate voltage is applied. Electrons from bP tunnel into HfO$_2$ layer and are trapped below the conduction level. (b) Band diagram of AHA charge storage layer with a charge density of $1.5 \times 10^{13} cm^{-2}$ in the HfO$_2$ layer without top gate voltage. Trapped charges can optically be excited and removed from the charge storage layer. (c),(d) The band alignment changes with the trapped charged density with (c) and without (d) the top gate applied. Long-dashed line, short-dashed line, and straight line refer to charge density of 1.5, 0.5 and 0 ($^{13}cm^{-2}$), respectively.

and plot in 4.5a. The initial programming of conductance of the bP-PPT channel requires 10s of ms gate pulse to saturate the charge density in the trapping layer as shown in Fig. 4.5b. Different gate voltages and pulse time are considered, which result in the different conductive states as shown in Fig. 4.3. The shorter pulses on the order of μs can be used to program the devices in smaller steps of conductance.

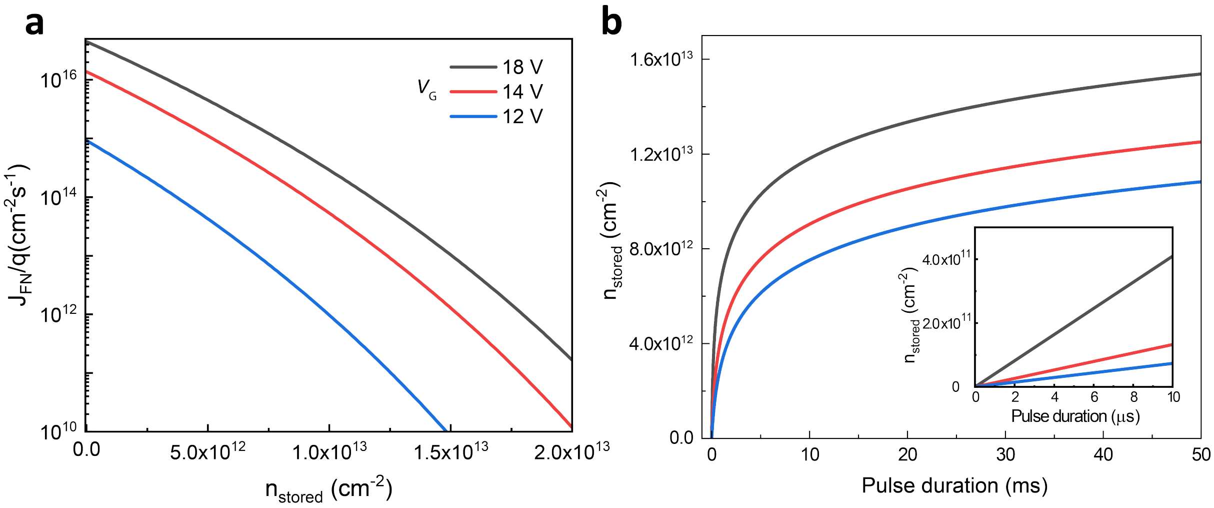


Figure 4.5: The electrical programming speed (a) The FN tunneling versus the trapped charge density in the storage layer with different top gate voltage V_G. (b) Different charge densities in the storage layers depending on the pulse duration. Inset: Zoomed-in plot of (b) with shorter pulses.

For optical programming, we used optical pulses of ms duration in our experiment. But based on the theoretical analysis, the programming speed can be increased with ns μs optical pulses, depending on the pulse intensity and photon energy. Mechanisms of detrapping of the stored charges in the HfO_2 layer include field-assisted tunneling, thermally excited charge tunneling, and optically excited charge tunneling as indicated in equation 4.8. Without the gate voltage, the tunneling coefficients for the first two mechanisms are negligible compared to the optically excited tunneling coefficient as shown in Fig. 4.6. Because the trapped charges have to overcome the large tunneling barrier with Φ_B, the charge tunneling process can be activated with optical illumination, which reduced the effectivd barrier to $\Phi_B - \hbar\omega$. The optically assisted tunneling process can be described as:

$$-\frac{dn_{stored}}{dt} \propto TC(\xi_{opt}, E_{built}(n_{stored})) * n_{opt} = TC(\xi_{opt}, E_{built}(n_{stored})) * (G(P_{opt}) * t) \tag{4.8}$$

where G is charge generation rate due to the optical illumination, which is assumed to be proportional to the optical power P_{opt}. To solve this nolinear differential equation, we assume the conductance of the bP channel is in the range of 30-80 μS, which corresponds to the stored charge density of 4 - 12 $\times 10^{12} cm^{-2}$. Also, we approximated the analytical function of TC with an exponential function and only considered the trap site with 1.25 eV below the conduction band, where the largest density of oxygen vacancies in HfO_2 exists ans reported with optical absorption spectra [112]. The resulting trapped charge density n_{stored} versus illumination time is plotted in Fig. 4.6b. The operation speed can be further reduced to the ns regime with higher photon energy and higher optical power. Considering our experimental conditions: the optical power on teh bP-PPT device is 12 μW, pulse width is varied from 1 ms to 200 ms, and photon evergy is 1.6 eV (780 nm). We use equation 4.8 to calculate the optical pulse energy required to change the device from state (n-1) to state #n, which is plotted in Fig. 4.6 c. The model shows a good agreement with the experimental results in Fig. 4.7 e.

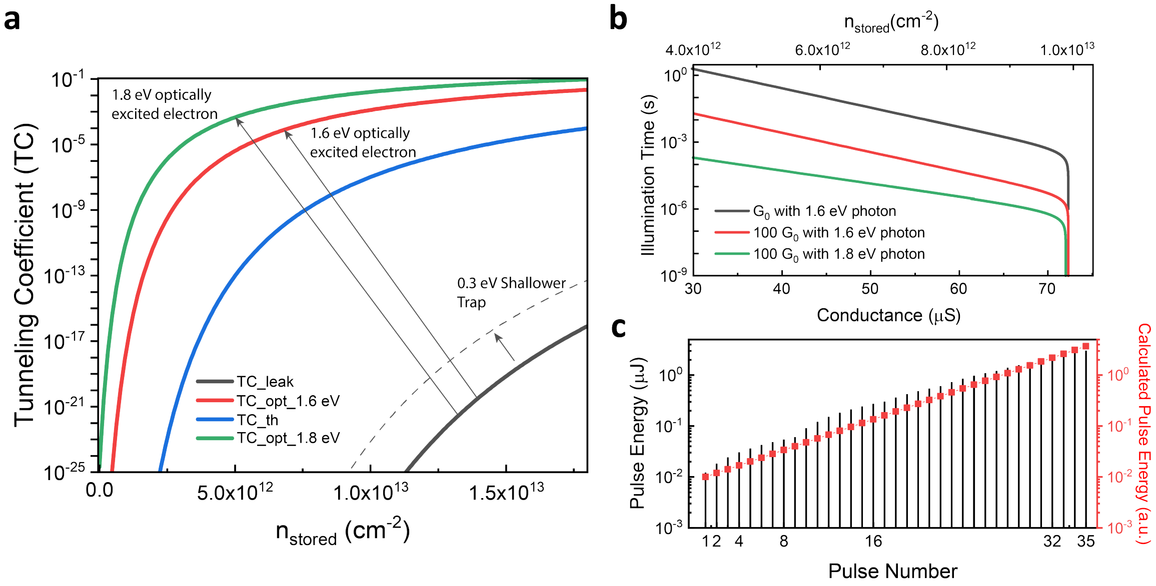


Figure 4.6: (a) Tunneling coefficients depending on the stored charge density in the storage layer. (b) The conductance of bP-channel changes depending on the optical power and illumination time. (c) The calculated and experimental optical energy depending on the pulse number n, that changes the device from state (n-1) to n.

4.3 Electrical and Optical programmability of bP-PPT

Figure 4.7 a illustrates the working principle of electrically programming the bP-PTTs by applying voltage pulses to the gate to enable charge tunneling from the bP to the HfO_2 layer. The device can first be reset with a depressive pulse (-18 V amplitude, 50 ms duration) to a fully-off state with low conductance. Afterward, it can be programmed by applying positive voltage pulses with amplitude in the range of 10-18 V and a fixed duration of 20 ms (18 V for state #0). By varying the pulse amplitude, the device can be programmed to states of more than 8 distinguishable and stable levels (equivalent to 3 bits) in its conductance when the bP channel is changed from p-type to n-type doping. The tunneling process and the resulting charge density can be modeled with the Fowler-Nordheim tunneling theory (see the previous section for detailed modeling methods). Figure 4.7 c and d show results of four representative states with a long retention time >2000 s (Fig. 4.7 c) and linear I-V characteristics (Fig. 4.7 d). The latter is important for its application in analog computing. Even higher precision can be achieved by programming the devices optically because optical pulses can directly excite the stored charges to remove them (Fig. 4.7 b), and the duration of optical pulses can be controlled more accurately than voltage pulses. We demonstrate optical programming of the bP-PPT devices using optical pulses in the wavelength of 780 nm, which provides sufficient energy to activate the stored charges to overcome the trapping potential (Fig. 4.7 b). Before programming, the bP-PPT is initialized electrically to state #0. Subsequently, it is illuminated with optical pulses with fixed average power (10 μW at the device) and varying duration so the pulse energy is varied between 10 nJ and 2 μJ. As shown in Fig. 4.7 e, these optical pulses program the bP-PPT to 36 states with different levels of conductance to represent 5 digital bits, a record-high number of levels achieved in charge storage devices. The programming process is accurate, arbitrary, and repeatable. The inset of Fig. 3.4e shows three adjacent levels that can be programmed repeatedly with high precision. The narrow bandgap of bP enables the bP-PPTs to be operated as broadband photodetectors that can detect optical signals from the near-infrared (NIR) to the mid-infrared (MIR) spectral range. Earlier studies have reported that a bP photodetector's responsivity is sensitive to the doping level and type of the bP channel [60]. In our bP-PPTs, since we can control the density of the stored charge to modulate the doping level and type of the bP channel, we can program their photoresponsivity in the same way as their electrical conductance. Figure 4.7 f shows the photoresponsivity of a bP-PPT measured in the wavelength range from 1.5 to 3.1 μm when the device is set to high and low conductance states (corresponding to states #0 and #35 in Fig. 4.7 e), respectively. Note that the low conductance state (state #35) has a high photoresponsivity due to the Burstein–Moss effect [66]. The unmeasured spectral range (1.8-2.6 μm) is due to the tunability gap of the light source (M-Square Firefly IR). The linearity of the devices' photoresponse is also verified for an incident optical power of up to 30 mW. Therefore, the bP-PPT has a programmable photoresponse in all the telecommunication bands (S, C, and L bands) and the mid-infrared range.

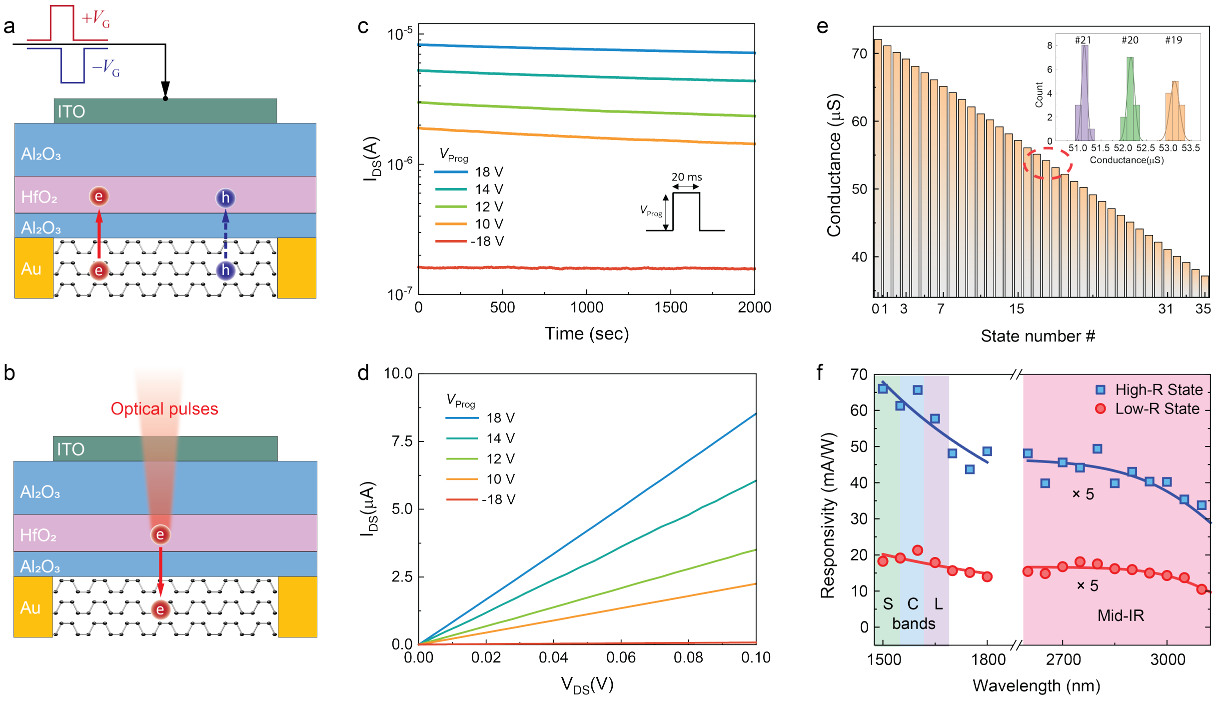


Figure 4.7: Schematic illustrations of the working principle of programming the bP-PPT using (a) electrical gate voltage pulses and (b) optical pulses. The red and blue circles represent the electrons and holes in the bP channel or HfO_2 trap sites, respectively. V_G is the applied top gate voltage. (c) The bP-PPT can be electrically programmed to 5 states of well-resolved conductance levels using pulses of different voltage amplitudes (V_{prog}) with the fixed pulse duration of 20 ms. Note that a negative repressive pulse (-18 V) resets the device to the lowest conductance. Inset: Illustration of the pulse shape of 20 ms in duration and V_{prog} in amplitude. (d) The I-V characteristics of the device at each programmed state, showing the linear conductance (g_{bP}). (e) The bP-PPT can be optically programmed using visible pulses to 36 levels in conductance. The inset shows the histogram of the well-separated conductance when the device is programmed repeatedly to three adjacent states indicated with the red dashed ellipse in the histogram. (f) The bP-PPT's photoresponsivity over the near-IR (including the telecom S, C, and L-bands) and the mid-IR ranges. The discontinued spectral region is due to the gap of the laser tunability. The bP-PPT's photoresponsivity can be programmed to two states when its conductance is set to high (state #0 in e) or low (state #35 in e). The solid lines are guides to the eye.

4.4 Multispectral Optoelectronic Edge Detection

The above results show that the bP-PPT devices can be programmed both electrically and optically. The programmed state is non-volatile and can be read out either electrically by measuring the device's conductance or optically by measuring its photoresponsivity. In both cases, the devices are operated in the linear regime and thus can be utilized for analog computing. Such a hybrid of multifunctional operation modes enables the utilization of a bP-PPT array to implement a mixed-mode optoelectronic neural network system. The same bP-PPT array can act as both the optical frontend to receive and preprocess optical images and an electrical processor with in-memory computing to post-process the images (Fig. 4.8 a). We first use the bP-PPT array to detect infrared optical images and preprocess them in the sensor[110, 111, 115, 116]. To prove the concept, we configure the bP-PPT array to perform edge detection of images by programming their photoresponsivity (R) to represent convolutional kernel matrices and receiving input images transmitted and encoded in the power (P_{in}) of telecom band optical signals. Measuring the photocurrent output $\mathbf{I}_{Ph} = \mathbf{R} \cdot \mathbf{P}_{in}$ in from the array corresponds to a multiply-accumulation (MAC) operation[117, 118, 119] on the input image with the kernel matrix stored in R. For edge detection, the photoresponsivity matrix R of a 2 × 2 bP-PPT array is optically programmed to binary values (Fig. 3.4f) and, after proper normalization, to represent kernel matrix $\left[\begin{smallmatrix} -1 & 1 \\ -1 & 1 \end{smallmatrix}\right]$ for right edge detection ($\left[\begin{smallmatrix} 1 & 1 \\ -1 & -1 \end{smallmatrix}\right]$ for top, $\left[\begin{smallmatrix} 1 & -1 \\ 1 & -1 \end{smallmatrix}\right]$ for left, and $\left[\begin{smallmatrix} -1 & -1 \\ 1 & 1 \end{smallmatrix}\right]$ for bottom edges)[120, 121, 37].

To demonstrate the broadband capability of the bP-PPT array, we encode three different 8-bit grayscale images (Fig. 4.8 b; top: handwritten digits; middle: a husky dog; bottom: a cameraman) using wavelengths in three telecom bands: 1510 nm in the S band, 1550 nm in the C band, and 1590 nm in the L band, respectively. The brightness of each pixel is encoded into the optical power using variable optical attenuators (VOA) and illuminated on the array. Each bP-PPT device is set to have a high responsivity of 60 mA/W to represent 1, or a low responsivity of 20 mA/W to represent -1 (Fig. 4.7 f). The measured photocurrents are normalized and offset to calculate the convolution. The convolved images without any further post-processing are shown in Fig. 4.8 c to f, for right, top, left, bottom edges, respectively. Figure 4.8 g shows the combination of all types of edges, resulting in a clear silhouette of each image. Thus, we demonstrate the bP-PPT array's application as an optical frontend capable of multi-spectral imaging reception and preprocessing.

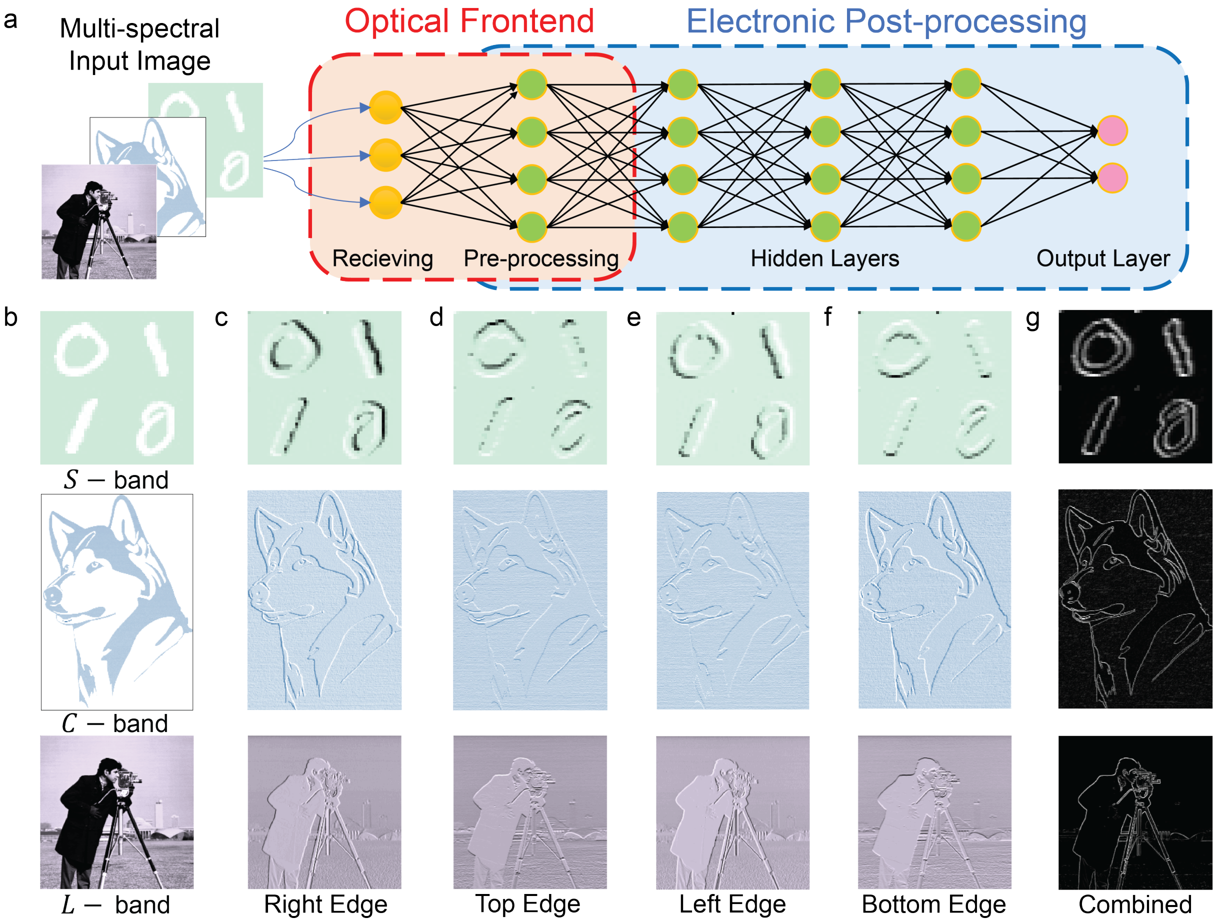


Figure 4.8: (a) The bP-PPT array receives images in multiple wavelength bands. The array's photoresponsivity matrix is programmed to represent the convolution kernel to directly preprocess the images in the optoelectronic domain (red dashed line box). The array's conductance matrix is then programmed to perform inference computation in the electrical domain (blue dashed line box). (b) The original input images encoded in the optical power transmitted in three different telecom bands. Top: handwritten digits (56 × 56 pixels, S-band); middle: a husky dog (312 × 222 pixels, C-band); bottom: a cameraman (256 × 256 pixels, L-band). (c)–(f) The resultant images after convolution with the right, top, left, and bottom edge kernels, respectively. (g) The final images combining all the edges.

4.5 BP-PPT Convolutional Neural Network

Besides its photoresponsivity, the conductance of the bP-PPT array can also be programmed to perform MAC operation by measuring the source-drain current $I_{DS} = V_{DS} * g_{bP}$, where g_{bP} is the conductance matrix of the array programmed to represent the weight matrix, V_{DS} is the source-drain voltages applied to the array as the input vector. An optoelectronic convolutional neural network (CNN) thus can be implemented with the array connected to the previous sensory devices, where the optical input image is detected and converted to electrical signals (Fig. 4.9 a, red dashed box). In Fig. 4.7 e, we have demonstrated precise programming of the bP-PPT to 36 discrete levels, ensuring high accuracy in weight training and inference calculation[122, 123]. Here, we use the 3 × 3 bP-PPT array to demonstrate a CNN that recognizes images of handwriting numbers "0" and "1" from the MNIST data set. The CNN consists of an input layer that captures a 28 × 28-pixel image, a convolution layer with two 3 × 3 kernels, an average pooling layer followed by an 8 × 2 fully connected (FC) layer (Fig. 4.9 a). The network is trained offline with 12,000 images of the training set with 100 epochs, delivering the final output scores that classify the input image to "0" or "1" with 99% accuracy. The trained network model is remotely programmed into the bP-PPT array by illuminating each pixel with the programming optical pulses. The kernel elements are discretized to accommodate the 36 discrete levels of the programmable states and used consistently in the experiment and simulation (Fig. 4.9 b). For example, the element value 2.00 in kernel 1 (K1), the largest element, is represented by setting a pixel of the bP-PPT to state #35 (in Fig. 4.7 e). By optically programming the 9 pixels of the array to the kernel elements, encoding the image pixels in the source-drain voltages, and measuring the source-drain current, the convolution calculation is executed on-chip to obtain the feature maps, followed by the average pooling and FC layers. The two output nodes from the FC layer are activated with the Softmax function and stored as scores to complete the classification task.

To verify the accuracy of the bP-PPT optoelectronic CNN, 100 randomly chosen images of handwritten numbers (48 of "0"s and 52 of "1"s) from the MNIST dataset were tested. The results are compared with the simulated results obtained from a computer. Note that this simulation is different from the first training with 99% accuracy due to the limited 36-level discreteness of the kernel element values. The bar graph in Fig. 4.9 c compares the experimental and the simulated output scores of the two labels "0" and "1" for 50 test cases, which have shown excellent agreement. The gray-tarnished bars in the experimental data

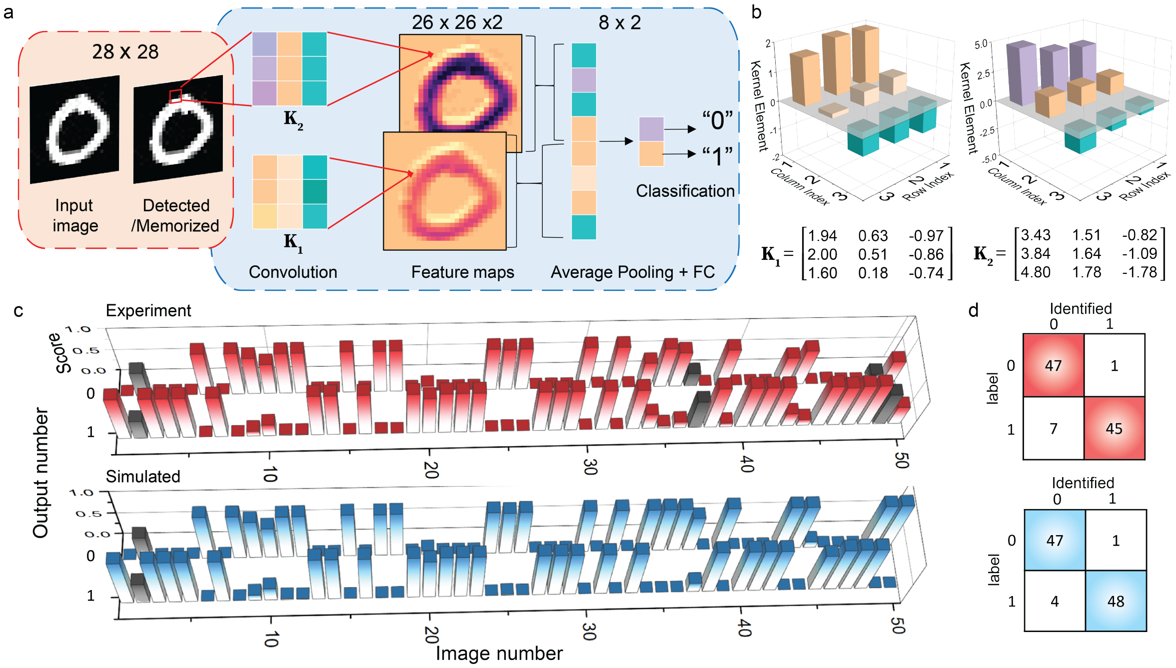


Figure 4.9: (a) CNN model for classifying handwriting numbers "0" and "1" from the MNIST dataset. The CNN consists of two convolution kernels, an average pooling layer, and a fully connected (FC) layer. The images captured by the bP-PPT array are further processed by the bP-PPT array in the electrical domain. (b) The 3×3 bP-PPT array is programmed with 5-bit precision to represent two kernels generated by offline training. (c) The experimental results (top, red) for image recognition using the bP-PPT array are compared with the simulation results (bottom, blue). Each bar is the score indicating the possibility of the CNN recognizing an image in the MNIST image library. The incorrectly recognized cases are in gray color. (d) The experimental and simulated confusion table for 100 images from the MNIST dataset. Colored diagonal elements in the table indicate the correctly identified cases.

highlight the incorrectly recognized cases. The experimental and the simulation results are summarized in the confusion table in Fig. 4.9 d. The bP-PPT array-based CNN reached an accuracy of 92%, comparable with the simulated results (95%).

4.6 Summary and Outlook

To summarize, we have demonstrated a phototransistor array based on bP (bP-PPT) that can be programmed electrically and optically by utilizing the stored charges in the gate dielectric stack that has a long retention time. Particularly, our device has a programming precision with a resolution higher than 5-bit, which is among the highest of non-volatile memory devices based on the charge trapping mechanism[124]. Leveraging its flexible functionality, we use the bP-PPT array to realize vision-sensory functions with in-memory computing. The sensors' programmable photoresponsivity enables in-sensor computing for edge detection on images that are optically encoded and transmitted over a broad infrared band. The same bP-PPT array can also be electrically programmed on the backend to implement a CNN to perform image recognition tasks. Although the demonstrated 5-bit programming precision of our devices is far less precise than that of digital computers, its application in analog in-sensor computing is more suitable for edge computing requiring low power consumption and low latency[116, 125, 126, 127]. With the electrostatic tunability of 2D materials, many researchers have demonstrated electrically/optically programmable conductance/persistent-photoconductance (PPC) using varied materials for electrical channels and charge storage layers as summarized in Table 4.1. On top of PPC, the programmable photoresponsivity (R) of the bP-PPT device enables instantaneous in-sensor image processing. The multifunctional operation in both optical and electrical domain together with the array geometry realizes CNN to recognize handwritten digits with 92% accuracy. The optical programmability sets over 5-bit of stable states and suggests parallel and remote programmability. Additionally, the demonstrated programmable photoresponsivity in the near-IR can be extended to a broader range of infrared and further improved by heterogeneous integration of bP with other 2D materials[128, 67, 129], or optimized for a specific spectral range by varying bP's thickness. It will allow multispectral image processing on edge devices, which can expedite many processes in industrial or biomedical applications[130, 131, 132, 133, 134, 135, 136, 137, 138]. Furthermore, recently reported centimeter-scale growth of bP suggests that it is promising to scale the bP-PPT array to an even larger array

Ref.	Channel	Storage	Programming method	Weight	Input	CNN Accuracy	# of stable states
This work	bP	AHA	Electrical & Visible light	g, R	E, O	92	36
[115]	bP	POx	UV light	g	E	90	N.D.
[140]	WSe_2	h-BN	Electrical	g, R	E, O	90	N.D.
[110]	MoS_2	AlOx	Electrical & Visible light	g	E	94	4
[141]	bP	AHA	Electrical	g	E	N.D.	2
[142]	bP	cPVP	Electrical	g	E	N.D.	5
[143]	MoS_2	Graphene	Electrical	g	E	N.D.	N.D.
[144]	InSe	Graphene	Electrical	g	E	N.D.	16
[111]	WSe_2	No memory	Electrical	R	O	99	N.D.

Table 4.1: Comparison of key features and performance of bP-PPT with those of prior works
* Input E: electrial, O:optical
* Weight: gfor conductance, R for responsivity
* N.D.: Not demonstrated
* AHA: $Al_2O_3/HfO_2/Al_2O_3$ stack

of megapixels[139]. Thus, the demonstrated multifunctional optoelectronic bP-PPT array, combined with parallel imaging and programming schemes, such as spatial light modulation and wavelength division multiplexing, can realize more complex deep neural networks for machine vision sensors distributed with edge computing.

Chapter 5

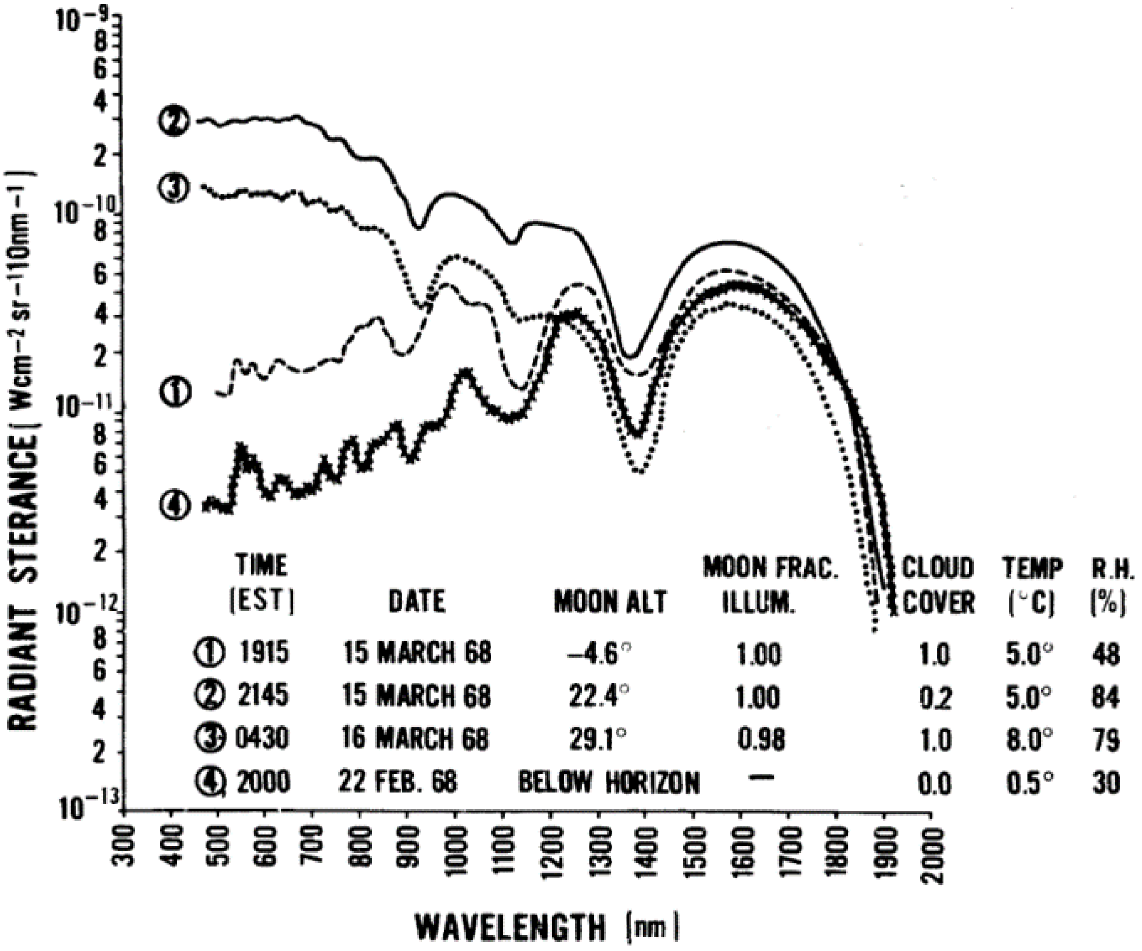
Heterojunction Phototransistor with High Gain

5.1 Motivation of Monolithic Night Vision Camera

5.1.1 Photon budget in Night time

Infrared detection helps us to see what human cannot see at night where visible photons are scarce. Figure 5.1 shows nightglow spectral radiance at the level of the ground with respect to moon's phases, measured by M. L. Vatsia et al [1] in the seventies. Night is brighter than people might think, since nightglow radiation is induced by radiative decay of atoms and molecules in the upper atmosphere, especially from the hydroxyl molecule (OH) produced in excited states through reaction between hydrogen and ozone. Hydroxyl molecule returns to a lower energy state by various processes including emission of radiation, resulting in the maximum radiance value received at ground level is in the Short Wave Infrared (SWIR) band, between 1.4 and 1.8 μm. Its level may fluctuate temporarily during the night, from one night to another and from one season to another, as well as spatially with the geographic location. Nevertheless, the nightglow radiance allows us to estimate photon numbers available for passive night vision with known imaging parameters.

We assume the objective has f/# of 1 and has a transmission of the lens T_{lens} of 50%, with an aperture of d of 1 cm and focal length f of 1 cm. The radiance is integrated between 0.8 to 1.6 μm to achieve a radiance



	TIME (EST)	DATE	MOON ALT	MOON FRAC. ILLUM.	CLOUD COVER	TEMP (°C)	R.H. (%)
①	1915	15 MARCH 68	−4.6°	1.00	1.0	5.0°	48
②	2145	15 MARCH 68	22.4°	1.00	0.2	5.0°	84
③	0430	16 MARCH 68	29.1°	0.98	1.0	8.0°	79
④	2000	22 FEB. 68	BELOW HORIZON	—	0.0	0.5°	30

Figure 5.1: Night-sky radiant sterance spectra recorded at Camp A. P. Hill. Reproduced with permision from [1]. Copyrights 1972.

R of

$$R = 2.31 \times \frac{10^9 W}{cm^2 Str}. \tag{5.1}$$

For an object at a distance L away, the solid angle Ω at the objective aperture is given by

$$\Omega = \frac{Area_{lens}}{L^2} = \frac{\pi d^2}{4L^2} \tag{5.2}$$

Assuming the image size of the object is same as that of a pixel size of p, then the object at a distance L with size D will be related to the lens parameters as

$$\frac{D}{L} = \frac{p}{f} \tag{5.3}$$

Hence, if the target object has 50% of reflectance (Refl), the optical power P_{opt} incident on each pixel becomes

$$P_{opt} = T_{lens} \times Refl \times R \times \Omega \times D^2 = T_{lens} \times Refl \times \frac{\pi}{4} \times P^2 \times (\frac{d}{f})^2 \times R \tag{5.4}$$

Interestingly, the power incident per pixel depends on the f/#, independent of the distance of the object from the objective. A larger pixel size or pitch will increase the incident optical power, at the cost of reduced Modulation Transfer Function (MTF, a measure of resolution of an imaging system). For the infrared sensor we are going to discuss with optical active area (p=80 μm), the incident power per pixel or single device becomes:

$$P_{opt} = 29.013 fW$$

Assuming an average photon energy of 0.8 eV ($1.28 \times 10^{19} J$), the photon flux becomes 226562 photons/sec. The estimated incidence power density is around 0.81 nW/cm^2 assuming 90% of transmission of optics.

5.1.2 Conventional Night Vision Intensifier

The past 40 years have seen minimal changes to the image intensifiers in night vision goggles. This is because GaAs photocathode image intensifiers lead the market due to their distinct advantages: >5000x light amplification, high SNR. However, GaAs photocathode image intensifiers have limitations in performance under cloudy moonless nights (restricted to 900 nm of IR wavelength), its heavy torque on neck (the weight is 600 g/monocle) and a narrow numerical aperture[145, 146, 147]. There has been work to address these limitations focusing on extending the absorption cutoff from GaAs (900 nm) to a region that is spectrally less sensitive to the phases of the moon (> 1000 nm) [148] and using semiconductor devices whose numerical aperture are not limited by the diameter of the multiplier tubes used in image intensifiers. Organic semiconductor devices have shown > 100x light amplification, but are still limited by the wavelength cutoff (1040 nm) [149]. Geiger-mode avalanche photodiodes are currently used in night vision, but due to requiring > 75V and signal processing to operate they are used as mounted cameras [**?**]. The devices that will be discussed in this paper use heterojunction phototransistors (HPTs), which use In 0.53 Ga 0.47 As absorbers to extend the wavelength cutoff to 1700nm. The best NIR to visible upconverter using InGaAs HPTs to date, are still under one-to-one photon amplification (59% p-p) [150] . This is in part due to the HPT used in [150] having low photosensitivity (50 A/W for 1.2 mW/cm^2 of input light) relative to what has been achieved in literature. To the best of our knowledge, the highest responsivities seen were between 3000-10k A/W for 3-6000 μW/cm^2 of input light , which is 100-260 x higher than the HPT used in [150] while still requiring 1M × more light than is available at night (1-10 nW/cm^2)[1]. In this work we show an HPT in linear mode able to upconvert 64% of 1550nm photons into visible photons. In avalanche mode, the HPT can upconvert 1,000,000% of 1550nm photons into visible photons at 6 nW/cm^2, which can be used at night light levels.

5.2 Device structure of Heterojunction Phototransistor

Bipolar junction transistors (BJT) have more than 70 years of rich history with its wide usages from logic electronics to electronic amplifiers. The earliest version of the transistor devices to be developed and produced was a homojunction transistor using Germanium, which was soon replaced by Silicon, as Germanium stops working above 75 °C [151] being impractical for most applications. Gradual performance improvements, especially with the developement of theory of the Heterojunction Bipolar Transistor (HBT) driven by Herbert Kroemer [152], which uses two or more different semiconductor materials with different bandgaps to enable high frequency and gain operation. His work earned him a Nobel prize in 2000. Though the theory was proposed as early as 1957, production of HBTs had to wait until 1977 for the advent of equipment that was capable of manufacturing it; first with MBE (Molecular Beam Epitaxy) and later with MOCVD (Metal-Organic Chemical Vapor Deposition). In this section we extend the HBTs to optoelectronic devices and explore InGaAs-based heterojunction bipolar phototransistor (HPT) as a shortwave infrared (SWIR) photodetector. By proper design of epitaxial layers, the HPTs demonstrates photoelectric gain as high as 10,000 electrons per photon at 1.2 V of bias.

5.2.1 Structure Design for High Gain Phototransistor

Extensive research has been conducted on $In_{0.53}Ga_{0.47}As$, which exhibits lattice matching with InP, making it a well-studied material for use in infrared photodetectors. Our device structure was grown using Molecular Beam Epitaxy (MBE) on a sulfur-doped n-type InP substrate with (001) orientation, as depicted in Fig. 5.2. The structure consists of 1.5 µm of n-type InP as the Emitter, doped at a concentration of 10^{18} cm^{-3}, 1 µm of p-type InGaAs as the Base, doped with 2.5×10^{16} cm^{-3}, and 500 nm of n-type InGaAs as the Collector, doped with 2.5×10^{18} cm^{-3}. Additionally, heavily doped n-type InGaAs is grown to facilitate Ohmic contact for device fabrication. The conversion of the epitaxially grown heterogeneous semiconductor film into pixelated photodetectors is achieved through nanofabrication techniques. The process commences with lithography and metal deposition of electrode pads measuring $100 \times 100\mu m^2$ on top of Heterojunction Phototransistors (HPTs). Ti/Au metalization, with a thickness of 20/150 nm, is evaporated and subsequently lifted off for establishing the Collector contact. Each detector is isolated by mesa isolation etching down to the InP emitter layer by wet (H_3PO_4-based solution) etches.

5.3 Optical and Electrical Properties of HPTs

Electrical properties of HPT can be understood with working principle of ordinary BJTs. Figure 5.3 illustrates the energy band diagram of InP/InGaAs HPT made of heterojunction NPN transistors in forward-active mode. Electrons injected from the emitter overcome the energy barrier qV_n by diffusion and enter the base. In general, the base width is small, compared to the minority carrier diffusion length, and therefore most of the electrons travel through the base due to their momentum to be collected by the collector. However, a small number of electrons are lost due to recombination in the emitter-base depletion region and in the base region. Now consider the holes in the bae which comprise the majority. They enter the emitter,

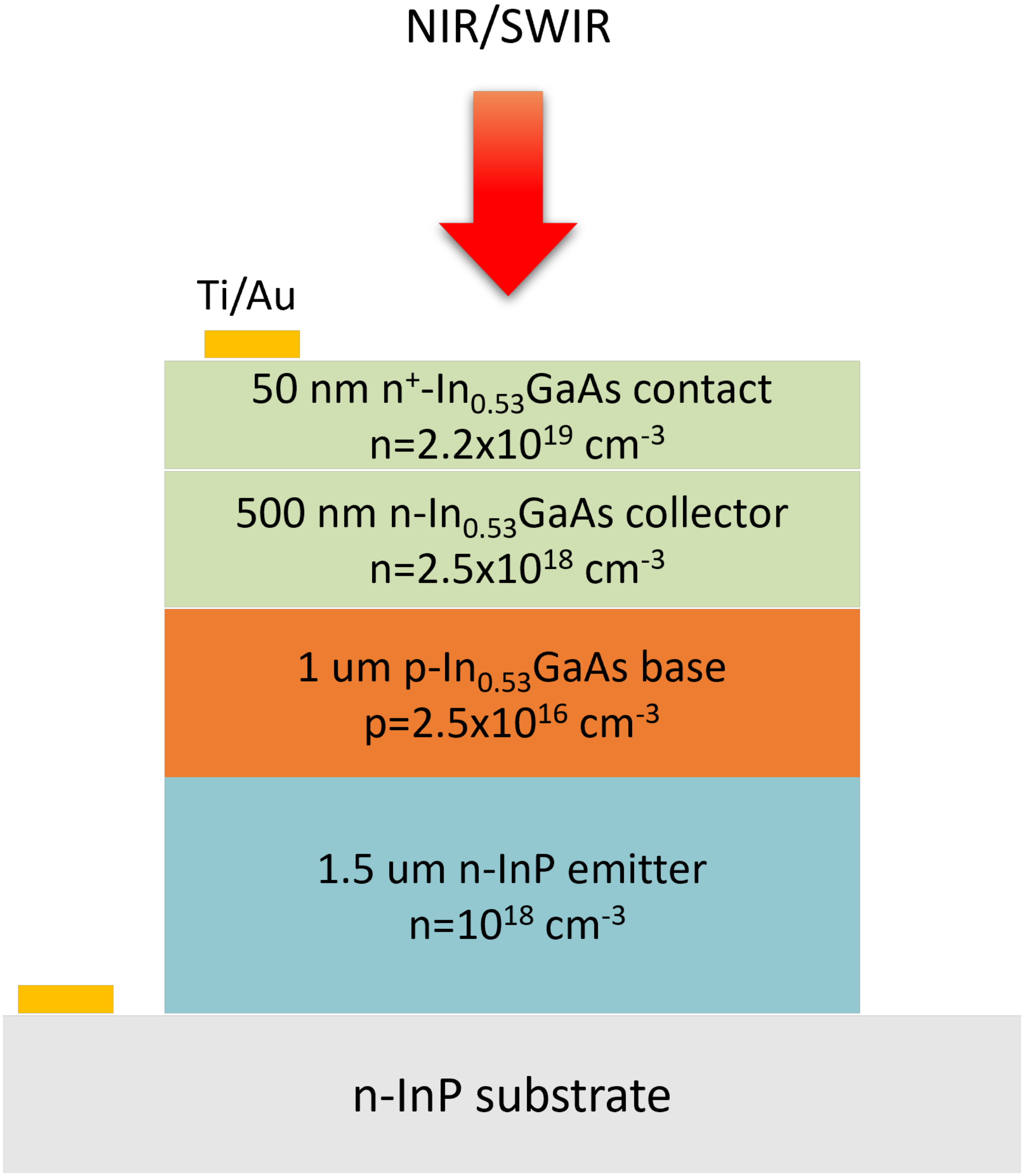


Figure 5.2: (a) Device structure of InGaAs-based HPT

overcoming the energy barrier qV_p. The large bandgap of InP is greater than the bandgap of InGaAs by ΔE_g which is the key to improved gain of HPT with large qV_p. The various currents in the transistor shown in Figure 5.3 (b) are as follows:

I_{En} : electron current from emitter to base

I_{Ep} : hole current from base to emitter

I_s : current due to electron / hole recombination within forward biased emitter-base space charged layer

I_r : current due to bulk recombination in base

I_E : emitter current = $I_{En} + I_{Ep} + I_s$

I_C : collector current = $I_{En} - I_r$

I_B : base current = $I_{Ep} + I_r + I_s$

The common emitter current gain is defined by:

$$\beta = \frac{I_c}{I_b} = \frac{I_n - I_r}{I_p + I_r + I_s} \tag{5.5}$$

Neglecting recombinations in depletion region and base (I_s, I_r), the maximum achievable β is [**?**]:

$$\beta_{max} = \frac{I_n}{I_p} = \frac{N_D^E}{N_A^B}\frac{v_n^B}{v_p^E} exp^{\Delta E_g / k_b T} \tag{5.6}$$

where N_D^E and N_A^B are doping levels of the emitter (Doners) and base (Acceptors), respectively.

v_n^B and v_p^E are the mean velocities of electrons from emitter to base and holes from base to emitter, respectively. k_b and T are Boltzmann constant and temperature.

Thus, the large bandgap difference (ΔE_g) of the heterojunction and the ratio of doping in Emitter and Base region play crucial role in increasing the base-to-collector current gain.

5.3.1 High Linear Optical Gain

The advantages in Common emitter gain of Heterojunction between Emitter and Base, and doping concentrations in BJTs were discussed. Conventional BJTs in electrical amplifiers are ternary terminal device, which means the three C-B-E components of a BJT should all be connected to external circuit. The voltage applied to B-E junction controls the flow of C-E currents, with the B-E current amplified by β in the

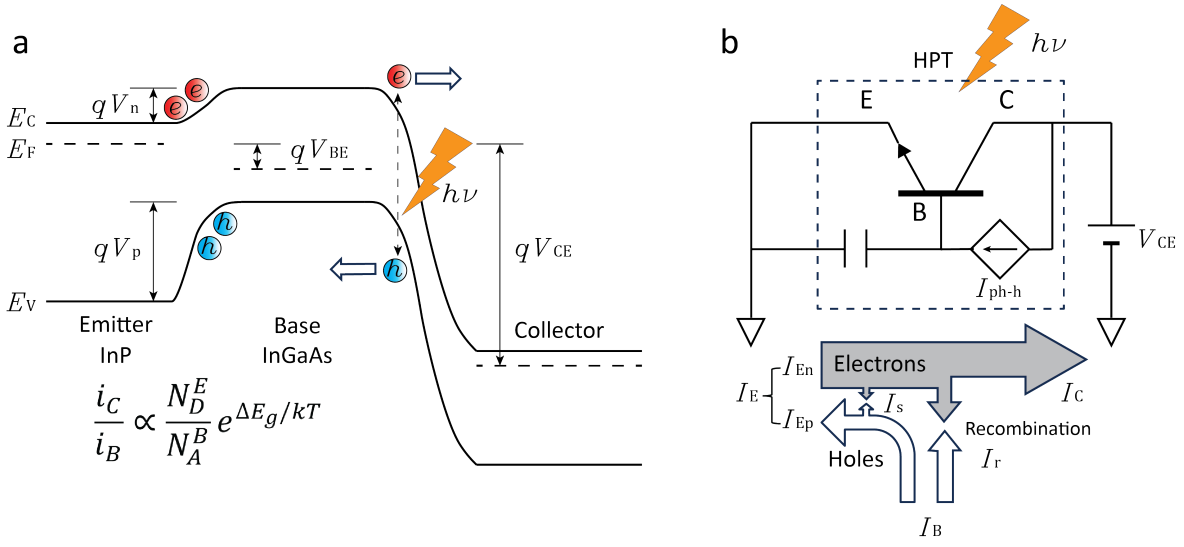


Figure 5.3: (a) Band structure of InP/InGaAs-based HPT with applied bias in Common Emitter topology. Infrared light with photon energy greater than the bandgap of $In_{.53}Ga_{.47}As$ (0.75 eV) can be absorbed either in the base, collector or base-collector depletion region. (b) Floating-base HPT operation mode with two terminal bias. The base current can be injected by optically generated e-h pairs

eqn.[5.6] and the large collector current is induced. It is important to note that applying current through C-E channel would not result in the voltage arising between B-E junction (below breakdown). It is nonreciprocal function; Voltage applied B-E junction controls the C-E current, not the other way. The current cannot control the voltage. Then how can we make use of the large gain of a BJT in a HPT as a photodetector, where the only two terminal is connected for operation? The key is holes optically generated in InGaAs layers, either in base or collector region. As shown in 5.3 (a), holes generated in Base and Collector region drift and diffuse to the interface between Base and Emitter, but the large barrier (qV_p) prevents the holes from flowing to Emitter. Assuming the recombination occurs slower than the accumulation, this built charges (correspondingly attracted electrons in the emitter region) act as a capacitor described in 5.3 (b) inducing bias voltage V_{BE} until the photo-generated hole currents (I_{ph-h}) matches the base injection current. This turns on the transistor even when the Base of the HPT is floated, and results in the large electron currents flowing from Emitter to Collector gained by β. The experimental results of our HPTs are shown in Figure 5.4. The transistor turns on at very low bias voltages under 100 mV, showing high Collector Output Resistance (6.8 MΩ, Early Voltage: 488 mV) with photocurrent gains as high as 3380 at 1 V of bias. To further characterize HPT at lower IR power regime, Noise Equivalent Power (NEP) measurement is performed. To suppress

Material	Mesa Diameter (μm)	Wavelength (μm)	Base Width (μm)	Base Doping (cm^{-3})	Optical Gain	Reference
InGaAs/InP	100	0.95-1.65	0.5	10^{18}	40	[153]
InGaAsP/InP	100	1.15	0.2-0.6	10^{19}	50-150	[154]
InGaAs/InP	20	1.15	2	5×10^{16}	600	[155]
InGaAsP/InP	147-346	1.15	0.5	1.5×10^{17}	1180	[156]
InGaAsP/InP	190	1.24	0.25	3.5×10^{17}	1000	[157]
InGaAs/InP	100	1.2-1.6	1.0	2.5×10^{16}	3000-10000	This Work

Table 5.1: Photo transistors for SWIR (0.9 - 1.6 μm Wavelength)

dark current noise while operation temperature is achievable with thermoelectric cooler, the measurement was done at 245 K with a cryo-probe station as described in Fig. 5.6 (a). Function generator generates a pulse train to modulate laser inputs as a sequence of on and off at certain frequencies. Photocurrent from the HPT is amplified as a voltage signal by preamplifier and measured by Lock-in Amplifier referenced at the function generator frequency with 1 s of integration time (1 Hz band). Figure 5.6(b) shows the measured photocurrent versus IR intensities at different modulation frequency of 100 Hz, 10 Hz, and 2 Hz. Horizontal dashed lines are noise floor measured by dark current spectral noise at corresponding frequencies indicated as colors. The linear photocurrent versus IR intensities over large dynamic range over 50 dB shows uniform responsivity of the HPT. Extrapolating the photocurrent versus IR intensities at low intensity regime to Noise floor, NEP is measured as 3 nW/cm^{2}Hz$^{1/2}$.

5.3.2 Enhanced sensitivity with combined Avalanche and transistor Gain

In Figure 5.4, distinct kinks beyond V_{CE} = 1.2 V are observed with nonlinear increase in the collector current. The steep rises of photocurrents are attributed to the avalanche carrier multiplication process in the reverse-biased Base-Collector region. Given the HPT is an NPN-bipolar junction device, positive collector-emitter voltage implies the reverse bias applied to base-collector region, which is effectively a PN-junction, made of InGaAs, operating in reverse bias. Typical InGaAs PN-junction shows avalanche breakdown around 20 50 V depending on doping profiles and geometries [ref]. As a result, the photocurrent generated by absorbing IR photons in the base-collector region experiences two gain mechanisms: transistor current gain h_{FE}, avalanche multiplication factor M. By solving for charge continuity, the current amplified by avalanche

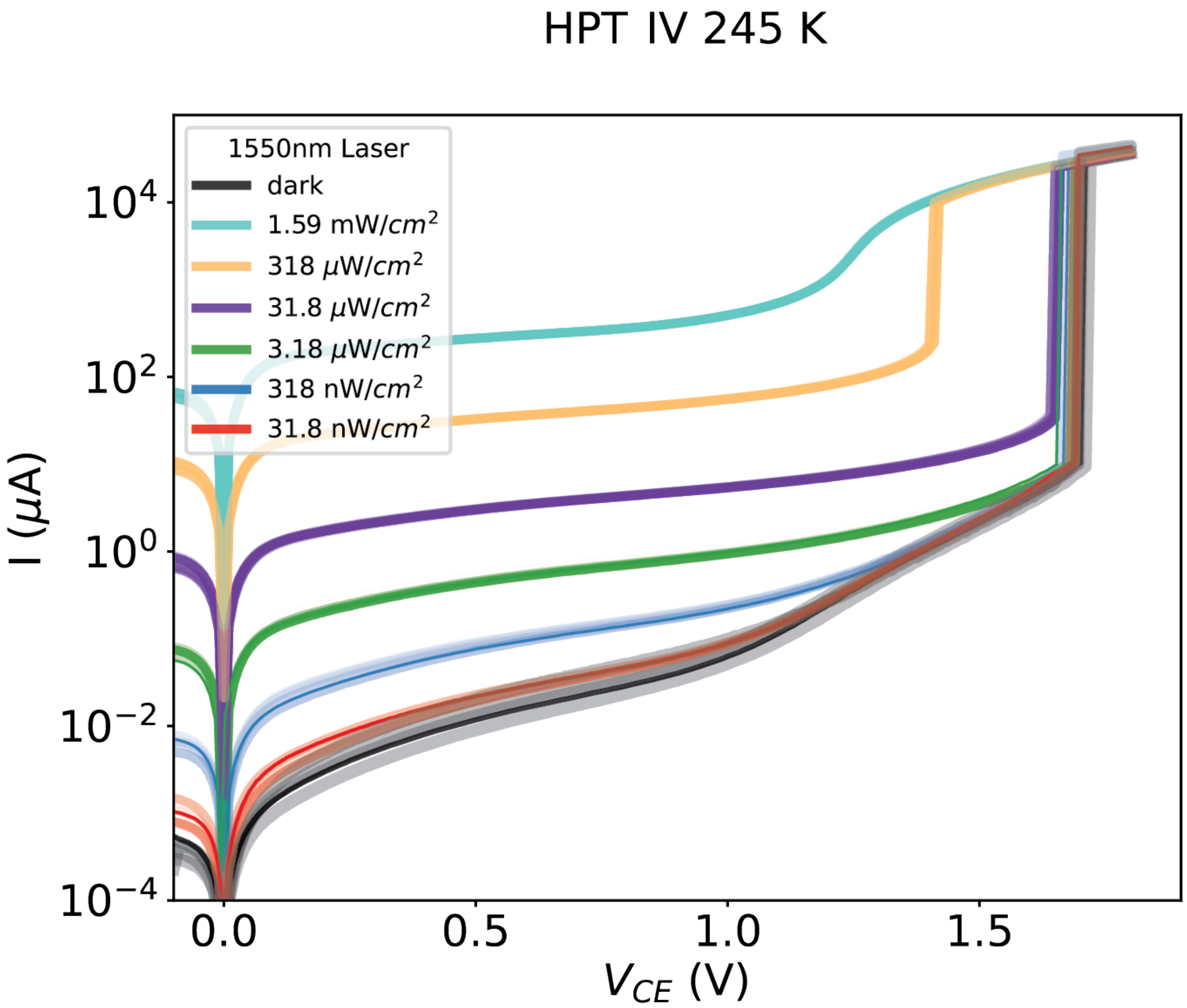


Figure 5.4: I-V characteristics of the HPT at different optical intensities at 1550 nm. The optical intensities shone here span more than 10^5 with the lowest intensities clearly distinguished from the dark current. Above 1.2 V of V_{CE}, large surge of currents were observed, which will be discussed in detail in the next section

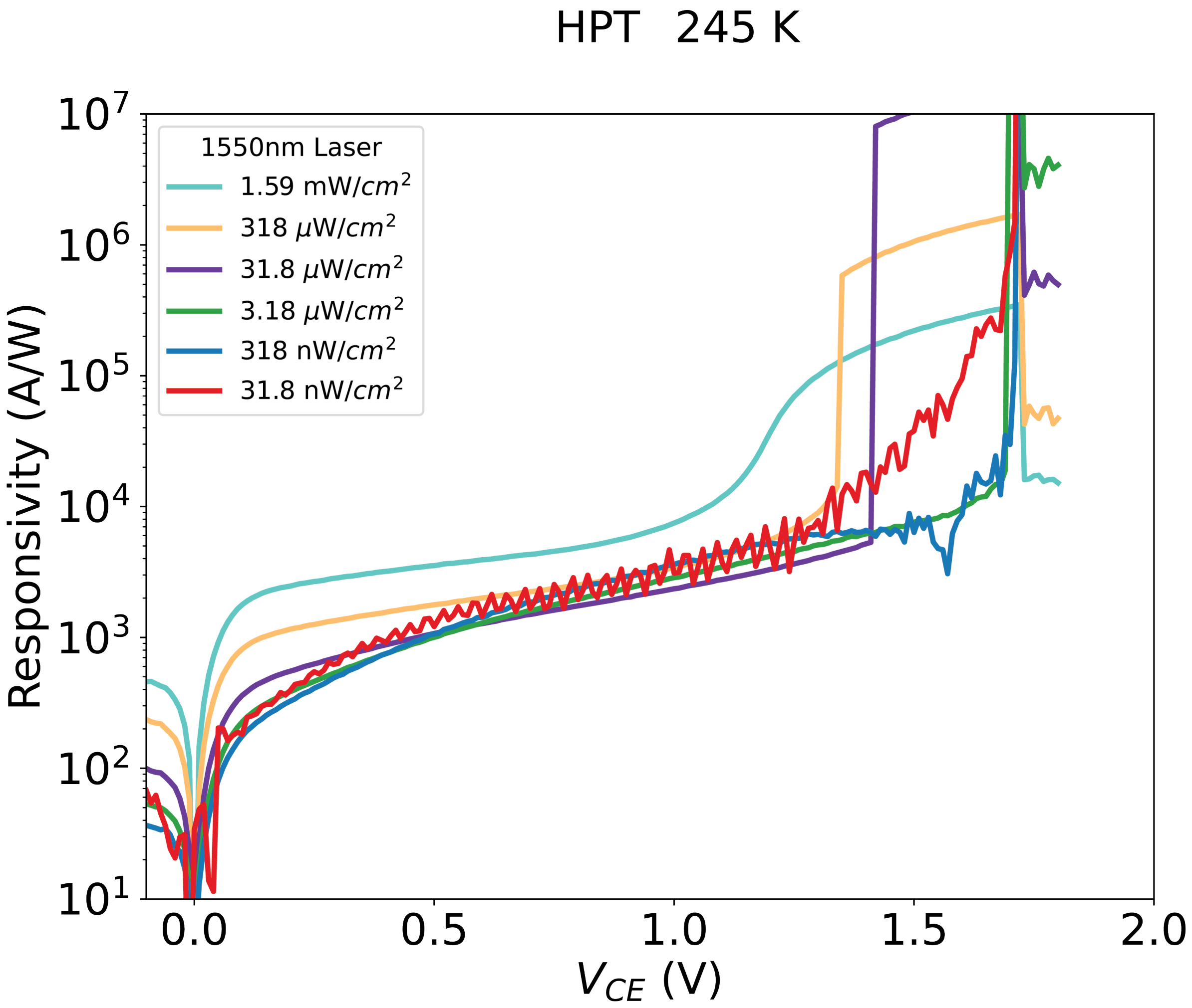


Figure 5.5: Responsivity can be obtained by subtracting dark current from collector currents, followed by dividing it with optical power, indicating the ratio of photogenerated charge to photon energy. The responsivity converges over large range of optical intensities at around 1 V of bias voltage. Between 1.2 V and 1.7 V, more than 1M responsivities are observed for intensities smaller than 320 μW/cm^2

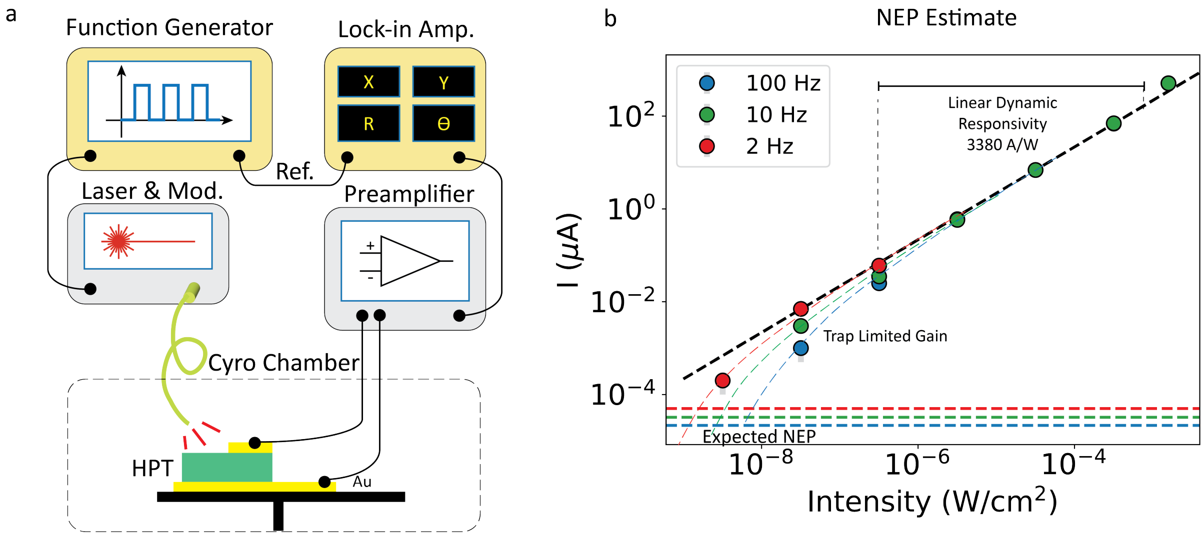


Figure 5.6: Noise Equivalent Power (NEP) measurement. (a) The NEP measurement setup. Function generator generates a pulse train to modulate laser inputs as a sequence of on and off at certain frequencies. Photocurrent from the HPT is amplified as a voltage signal by preamplifier and measured by Lock-in Amplifier referenced at the function generator frequency with 1 s of integration time (1 Hz band) (b) Measured photocurrent versus IR intensities at different modulation frequency of 100 Hz, 10 Hz, and 2 Hz. Horizontal dashed lines are noise floor measured by dark current spectral noise at corresponding frequencies indicated as colors. The linear photocurrent versus IR intensities over large dynamic range over 50 dB shows uniform responsivity of the HPT. Extrapolating the photocurrent versus IR intensities at low intensity regime to Noise floor, NEP is measured as 3 nW/cm^{2}Hz$^{1/2}$

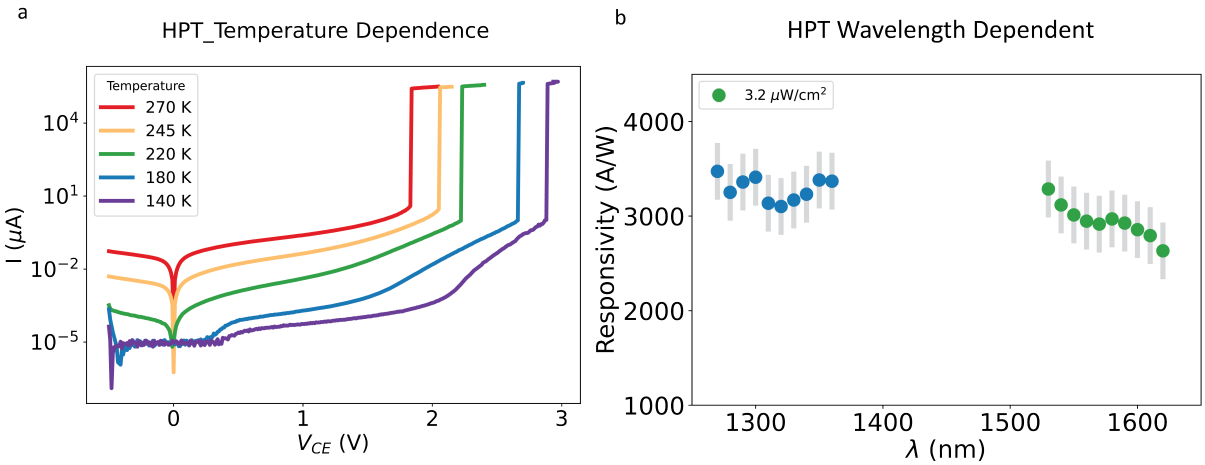


Figure 5.7: HPT IV curves of dark current at different temperatures and responsivity over broad wavelength regime (a)HPT IV curves of dark current at different temperatures. As opposed to the case of APD, the transition voltage increase with lowering temperatures. (b) The responsivity of HPT is maintained large over 3000 A/W in broad IR band from 1260 to 1630 nm.

HPT can be expressed as [ref]:

$$i_c = \frac{M(1+h_{FE}}{1-(M-1)h_{FE}}(i_{ph}+i_{dark}) \tag{5.7}$$

, where $M = [1-(V_{CB}/V_a)^n]^{-1}$ is avalanche multiplication factor and h_{FE} is HPT electrical gain. V_{CB} and V_a are bias voltage between collector and base, and avalanche breakdown voltage, respectively. n is a fitting factor typically ranging from 2 5. $i_{ph} = P_0\eta q/h\nu$ is primary photocurrent generated by IR photon (with power of P_0, quantum efficiency of η) without gain and i_{dark} is dark current generated by thermal excitation. Eqn. 5.7 suggests that introduction of avalanche multiplication factor slightly over unity, combined with the large transistor gain, leads to the collector current divergence much below the avalanche breakdown voltages. Note that in Fig. 5.4 the current divergence occurs below 1.7 V much smaller than InGaAs avalanche breakdown voltages (more than 20 V). This transition occurs at lower bias when weak IR light is shone. The apparent gain from the small IR intensities below 32 W/cm^2 exceeds 1 million, meaning a single IR photon generate a million electrons flowing across the HPT circuits, as shown in the Fig. 5.5 at 1.68 V. The difference between our Avalanche HPT and conventional Avalanche photodiode (APD) is characterized by their temperature dependent transition voltages, as well as the large difference in absolute values of bias voltages (below 2 V and more than 20 V). As indicated in Fig. 5.7 (a), the voltages of transition for HPTs increase with lowering the temperatures, as opposed to the case of APD, where the avalanche breakdown voltages decrease with lowering the temperatures, confirming the distinct mechanisms of Avalanche HPTs, as described in Eqn. 5.7.

5.4 Direct driving LEDs for Infrared Imager

The output impedance of the HPT is high enough to drive a serially connected LED, as usual transistors operating at common emitter mode. To be specific, the amplified current of the transistor is modulated by 1) Gate voltage for MOSFETs and 2) Base current for BJTs. Our HPT's output current is modulated by photocurrent excited by IR photon absorption and injected to base region, as discussed in the last section. Here, we will use the HPT to drive LEDs to complete a system where IR photon is converted to visible photon. Before combining the two devices, LEDs' characterization is performed.

5.4.1 LED Fabrication

Figure 5.8 explains the process flow for fabrication of LED array. The LED structure is devised from III-V based red LED developed in University of Illinois [83], with P-I-N structure where P and N type semiconductors are made of GaAs, and I region is made of InGaP where radiative recombination occurs to emit red lights. To pixelate the device, hard mask using SiOx grown by PECVD (Plasma-Enhanced Chemical Vapor Deposition) was used to define mesa of LEDs. The mesa formation is performed by ICP (Inductively coupled plasma) chlorine-based etching method. Since the byproduct of $InCl_2$ is nonvolatile, without proper chemistry selection and temperature condition for the etching process, etching of III-V materials easily fails with lots of bumps from self-masking of the nonvolatile byproducts, resulting in non-radiative carrier recombination, as shown in Fig. 5.8 bottom left. Adding CH_4 and Ar to BCl and Cl based chemical etching environment at an elevated temperature of 100 C solves the problem, producing smooth side wall of LED devices (Fig. 5.8) bottom right [158]. After the ICP-etching (also called dry-etching), SiOx are removed by wet-etching using HF-based solution, and the whole mesa structure is passivated by PECVD again to separate the P and N electrodes of the LED. By deposition of metal for electrodes with Ohmic contact with GaAs, the LED device fabrication is complete. Their images in an array and light emission is captured in a SEM image and an Optical Microscope image, in Fig. 5.8 middel row.

5.4.2 LED Characterization

The emissivity (measured in W/cm^2) of the LEDs are characterized using photodetectors and source-measurement-unit (SMU) to measure optical power emitted from the surface of the LED at specific Current density. Since an LED is a diode, or PN-junction, it also shows characteristic I-V curve with threshold voltage to turn the current on. Also, the light coming out of the LED shows characteristic J-L curve with threshold current density where it starts to emit a meaningful power of light. The measurement setup is as shown in Fig. 5.9. The emission profile from the surface of the LED is assumed to have Lambertian surface. The power collected by an objective lens with NA=0.4 is integration of Lambertian profile within the solid angle captured by the objective lens, which is 12 %. The power of the LED light collected by the objective lens is measured by Si-based photodetector with aperture size bigger than collimated LED light width down to 5 nW, and associated with current density. Below the power, APD was used for more accurate light emission from the

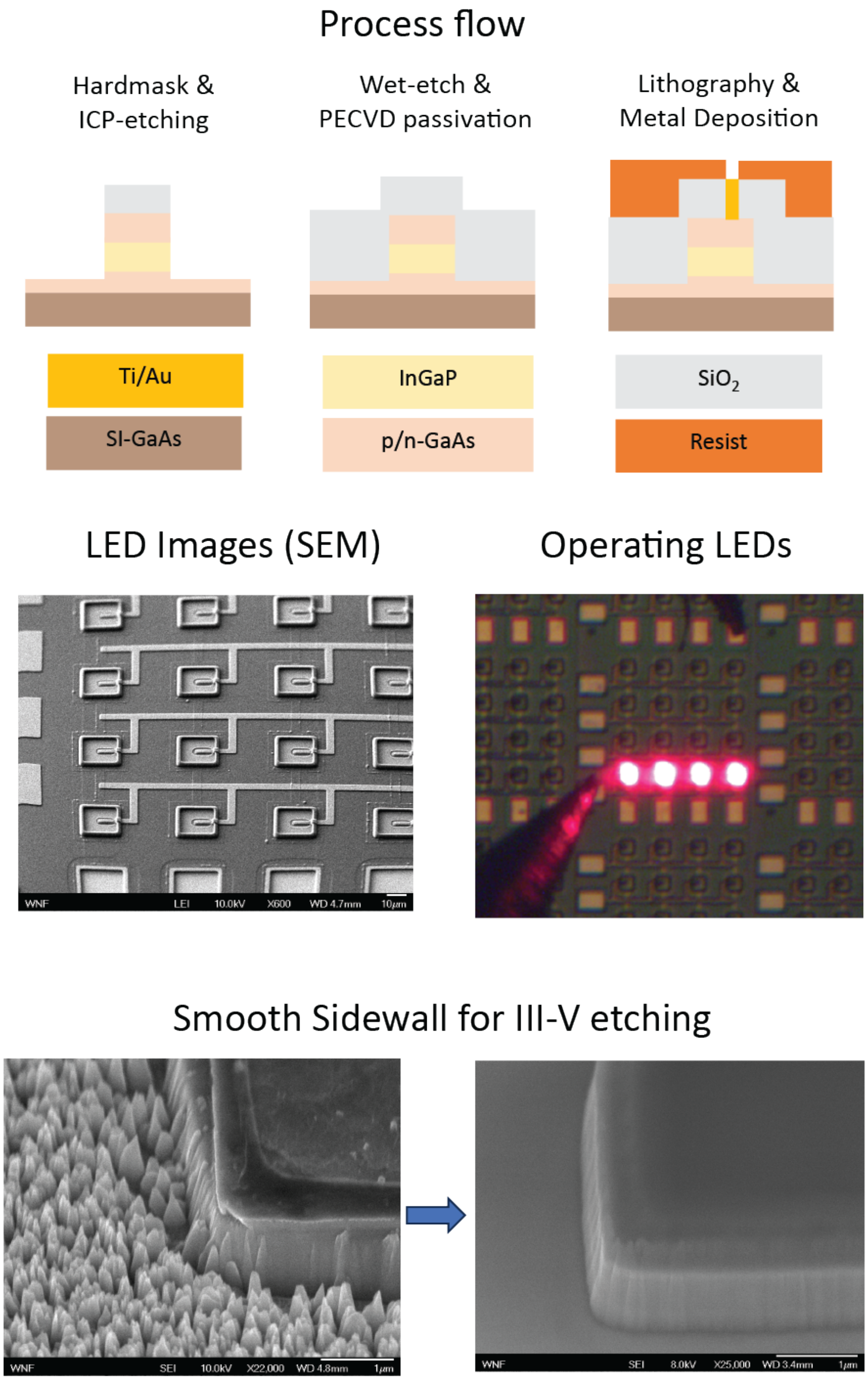


Figure 5.8: Fabrication process for LED mesa formation. (Top) Using hard mask and ICP-etching method, LED pixels are formed. Passivation and metalization follows to finish the LED device. (Middle) SEM and OM image for the LED array. Bright red light is emitted. (Bottom) The sidewall and etching quality depending on the ICP etching recipe. By selecting proper chemistry, smooth sidewall is obtained (right).

LED with calibration of APD counts per second (cps) to power of the light. Given the surface area of the LED is $100\mu m \times 100\mu m$, the threshold current density of LED is measured at 400 μA/cm^2. As shown in Fig. 5.9, the emissivity of the LED is measured as 10^5 nW/cm^2 at current density of 1.25×10^5 μW/cm^2. Since the photon energy at 620 nm (Red emission) is 2.0 eV, the external quantum efficiency (EQE) of the LED is estimated as 0.04 %.

5.4.3 HPT and LED Serial Connection

Transistors' high output impedance serves a great advantage to drive an arbitrary load. Our HPT, whose I-V characteristics shows large input impedance, can indeed drive an LED when they are serially connected, as shown in Fig. 5.10 (a). The I-V characteristics of the HPT+LED system illustrate the IR intensity-dependent photocurrent generated by the HPT, flowing through the LED, which can be directly translated to visible photon emission from the LED. The characterization of LED has been done with LIV measurement in the last section, having external quantum efficiency (EQE) of 0.04 %. The red dashed line in Fig. 5.10 (a) is the bias fixed for linear operation where one IR photon is converted to 1.6 red photon emission. In this mode, when the IR light is modulated as a pulse train of 100-Hz square-wave (pink pulses in the top subplot of Fig. 5.10 (b)), the photocurrent follows the optical pulses with rise time of 340 μs, fast enough for real time refresh rate of 60 Hz for human recognition. The photocurrent is high enough to turn the LED on over its threshold current density of 400 μA/cm2 (Supplementary Information), at as low IR intensity as 3.2 μW/cm^2. In large optical intensity range from 32 to 3200 μW/cm^2, IR-to-Red photon conversion ratio of 1 to 1.6 was maintained. Below it, degradation of the external quantum efficiency of LEDs limits the red photon output, not the HPT's photon-to-electron conversion, which were maintained over 3000 at as low IR intensity as 320 nW/cm^2. Operating in avalanche mode conveys its unique advantages when the IR light becomes even dimmer. The HPT + LED system at around 3.8 V, which corresponds to the transition edge voltage for dark condition, indicated as blue dashed line in Fig. 5.11 (a). When the bias is fixed slightly below the transition edge (let's call it arming voltage V_{arm}), the HPT + LED system becomes extremely sensitive to the IR input and the photogenerated current results in large collector current divergence from our avalanche HPT, limited by the LED current density, which is in the order of 1 A/cm^2 to exhibit bright red-light emission as high as 10^5 nW/cm^2 in the dim IR light of 6 nW/cm^2. The avalanche current is

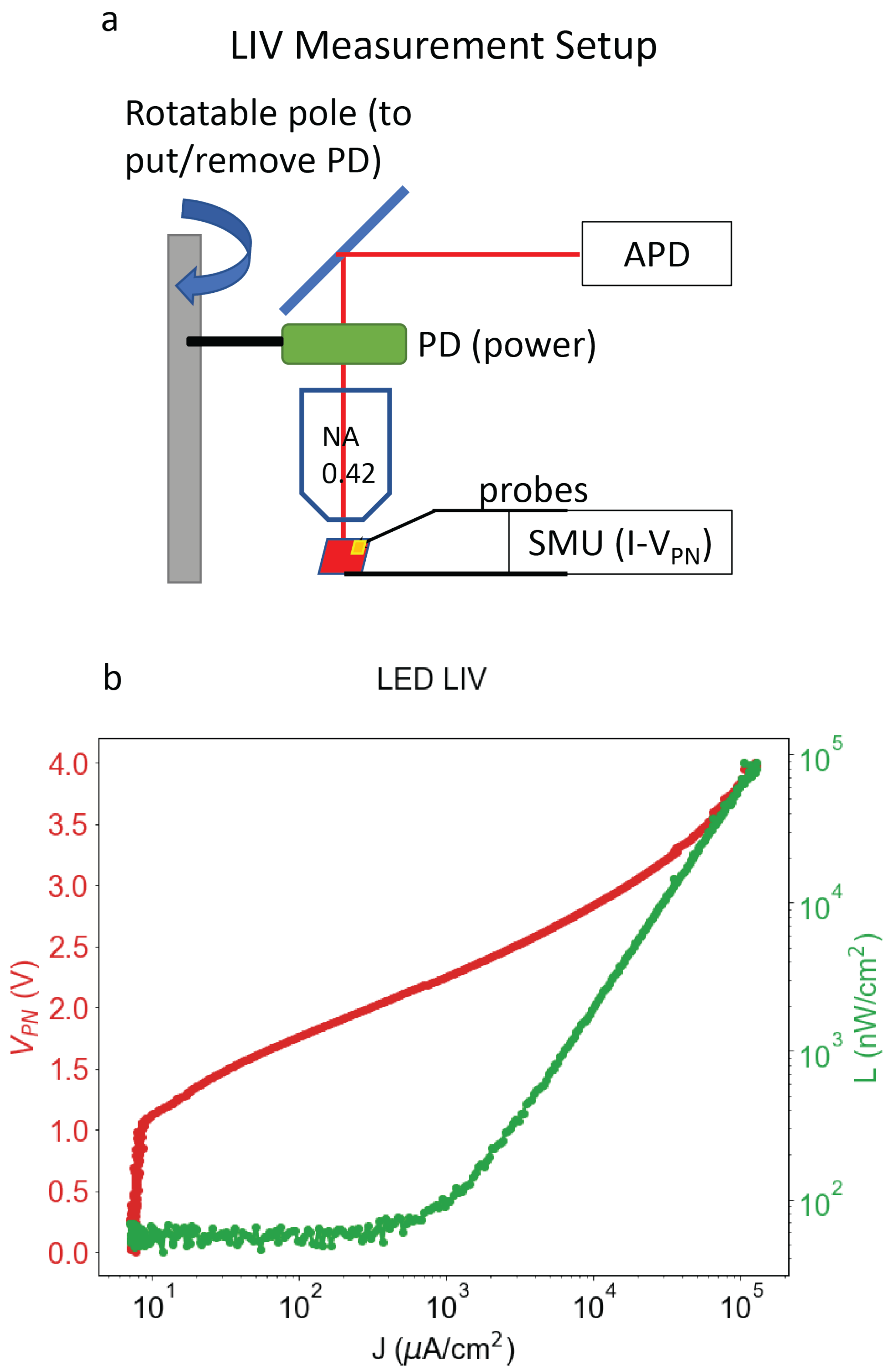


Figure 5.9: LIV characteristics of the LED. (a) Measurement setup using Photodetector and APD-single photon counter, and SMU for I-V to bias the LED. An objective lens is used for light collection. (b) LIV characteristics. I-V characteristics on left-y axis and L-I characteristics on right-y axis. Current density J instead of I is plotted.

self-sustaining once it initiates even after the IR light turns off. Thus, it needs to actively be quenched by reducing the bias voltage below the self-sustaining voltage (V_{quen}), and then biased again at arming voltage for the HPT to be prepared for the next detection. The sequence of arming and quenching of the HPT using active bias is indicated in Fig. 5.11 (b) middle subplot with green squares. Input IR pulses with intensity of 6 nW/cm2 are shone at half-frequency of the bias sequence with 10 % duty cycle (Fig. 5.11 (b) top subplot) to demonstrate the dynamic large current driven by the avalanche HPT. As illustrated in Fig. 5.11 (b), when the HPTs are ready for detection at the arming voltage indicated as green area, coincided with a dim IR incidence, the photocurrent large enough to drive the serially connected LED can flow through the system. This simple serial connection of HPTs and LEDs, driven by a bias voltage as being two-terminal system, becomes an easy solution for scaling up to full two-dimensional screen. Figure 5.12 (b) demonstrates pixelated LEDs lit up by serially connected to HPTs on breadboards. When both HPTs and LEDs are pixelated in 5 × 5, illuminating an IR image with 'I' letter shape on the HPTs, the LED array displays the letter 'I' as in Figure 5.12 (b) bottom.

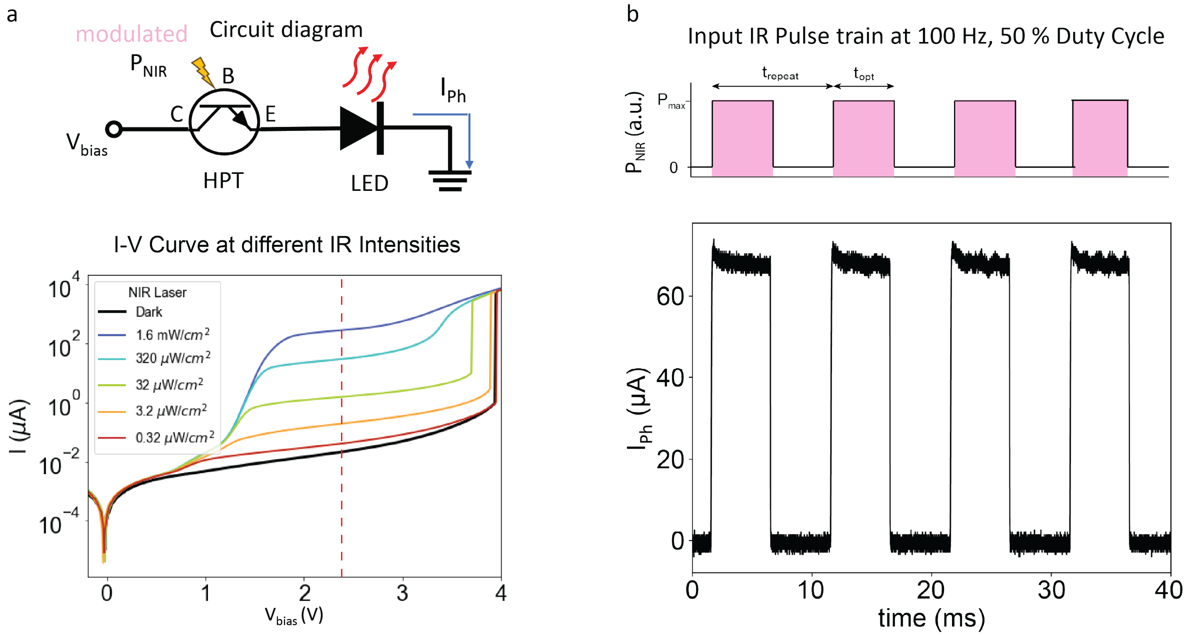


Figure 5.10: (a) Circuit Diagram for HPT + LED serial connection for IR-to-Vis photon conversion, top. The I-V characteristics of the system with bias voltage with respect to different optical intensities, bottom. (b) The IR input is modulated at 100 Hz with 50 % duty cycle as indicated as pink squared on top. The photocurrent responding to the modulated IR is measured at a fixed bias as red-dashed line in (a).

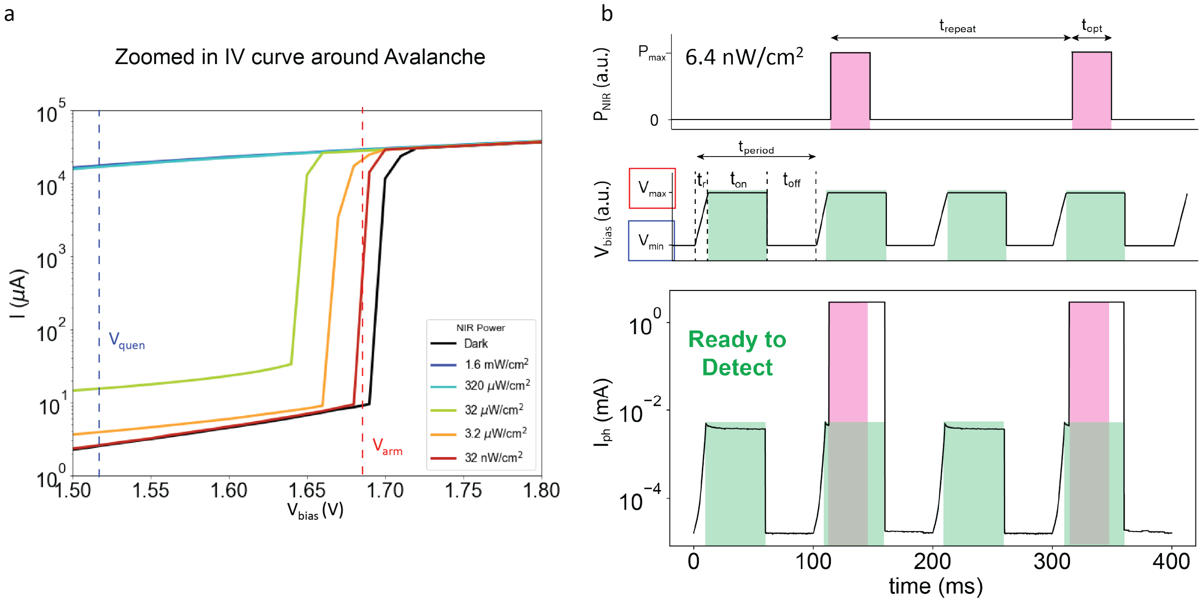


Figure 5.11: IR-Vis photon conversion at small IR intensity with avalanche mode. (a) The zoomed in IV curve around the transition edge. Red and blue dashed lines correspond to the arming voltage and quenching voltage, respectively. (b) (top) Modulated IR input pulse with half-frequency of the arming and quenching cycle. (middle) The bias voltage is actively modulated for HPT to be ready to detect low-intensity-IR at the arming voltage and to be quenched at the queching voltage. The green square corresponds to time zone when the HPT is prepared for detection. (bottom) The photocurrent response with the active biasing when low-intentisy IR inputs are coincided.

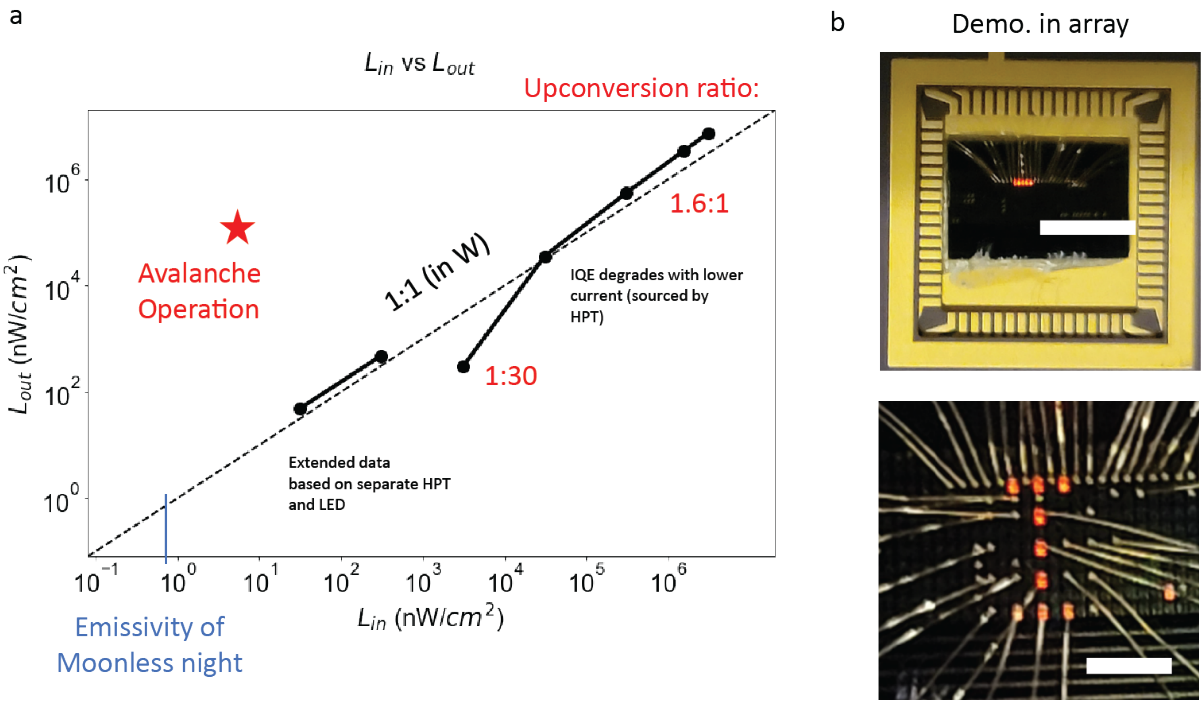


Figure 5.12: (a) L_{out} emissivity from LED versus L_{in} input IR intensity. Dashed line corresponds to 1:1 in power ratio. At large intensities 1.6 W IR photons generate 1 W of Red photons, resulting in 64 % upconversion. (b) Demonstration of HPT + LED array in 1-D (top) and 2-D (bottom). The scale bar is 5 mm for top, and 1 mm for bottom.

5.5 Summary and Outlook

To summarize, we have demonstrated a Heterojunction Phototransistor (HPT) with high gains in both linear and avalanche regime. The transistor gain for the linear mode has been optimized for high gain while maintaining decent low dark current at 245 K available with solid state thermoelectric coolers, by engineering heterogeneous material combination, dopants selection and doping level control, and diffusion length increase with high quality III-V crystals. The optimized HPT exhibits > 3000 p-to-e conversion, or photoelectric gain in linear regime in a wide range of IR intensities over 50 dB, also with broad wavelengths ranging from 1260 nm to 1630 nm. With the HPTs' high gain, even at small quantum efficiency of LEDs, the HPT + LED system can convert one IR photon to 1.6 red photon down to 3 μW/cm^2 of input IR intensity. There have been many research efforts in IR-to-visible light conversion, tabulated with their specifications in Table 5.2. Aside from high linear gain of HPT in broad IR range, avalanche mode of the HPT can drive an LED to emit red emission 10,000 times brighter than IR input intensity of 6 nW/cm^2. This level of IR intensity is 1000 × improvement from [149] in Table 5.2. Although at the expense of reduced time resolution to arm and quench the HPTs, the avalanche mode of HPTs offers high sensitivity of IR photon, expected to detect even lower intensities dimmer than 0.1 nW/cm^2 when the bias control can be done more precisely with stable circuitry, which can be truly intensify IR images in moonless dark night (for the emissivity of an objects of dark night, see Supplementary Information). Moreover, the operation voltage for HPT + LED system is less than 4 V, even at 2.3 V for linear mode, which is a big advantage for AA battery operation with long lifetime. All in all, the carefully engineered HPT has a promising application for night vision intensifier with the highest responsivity over 3400 A/W resulting in 64 % of 1550-nm IR photon to 620-nm visible photon conversion within large dynamic range of IR intensity from 3.2 μW/cm2 to 3.2 mW/cm^2. The high responsivity over broadband of Near IR beyond 1 μm to 1.6 μm increases the available photons for night vision intensifiers. In avalanche mode, even 1000 times dimmer IR light can be converted to red lights, corresponding to 1,000,000 % of 1550-nm IR photon to 620-nm red photon conversion, demonstrated for the first time. With LEDs with better quantum efficiency, the IR-to-Visible light upconversion using the HPTs is promising. In addition, the two-terminal operation of HPT + LED is beneficial for large scale integration for near eye display.

Ref.	PD (or HPT)	LED EQE	NIR-to-Vis (wavelength)	NIR-to-Vis (ph-to-ph)	NIR Intensity	Operating Voltage
[150]	50 A/W	1.5 %	1550 nm -> 564 nm	1.67:1	1.5 mW/cm^2	6-21 V
[159]		3.6 %	810 nm -> 590 nm	100:1.5	2 mW/cm^2	2 V
[160]	20 A/W	x	1550 nm -> 870 nm	50:7	0.74 mW/cm^2	3 V
[153]	100-1k A/W	x	1550 nm	x	0.1 μW	3 V
[149]	1000 A/W	1 %	1040 nm -> 580 nm	1:10	3 μW/cm^2	10-15 V
This work	3k - 1M A/W	0.04 %	1260-1630 nm -> 620 nm	1.56:1	6 nW/cm^2	2.3-4 V

Table 5.2: Demonstrated metrics for IR-to-Visible light upconversion

Chapter 6

Conclusion

The exploration and engineering of semiconductor materials have emerged as pivotal pillars in propelling the frontiers of infrared (IR) technology. The distinctive properties and manipulable characteristics of semiconductor materials serve as a keystone in the evolution of IR technology, revolutionizing diverse fields such as sensing, communication, imaging, and beyond.

6.1 Summary

Throughout this thesis, we have delved into the intricate world of semiconductor materials and their multifaceted properties, elucidating their indispensable role in advancing IR technology. The tailored manipulation of semiconductor properties, from bandgap engineering to charge carrier doping, has empowered the precise control and optimization of materials for specific IR applications. Moreover, using two-dimensional materials, optical properties are modulated with simple field effect devices. It is applied to an optical modulator in integrated silicon photonics, extending possibility of optical communication to mid-IR regime. Also, with dielectric layers engineered according to their bandgap, programmable phototransistors were demonstrated to exhibit multiple functions, and deployed to in-sensor computing. As well as two-dimensional materials, III-V compounds semiconductors have promising application in infrared technologies, as we demonstrated it with high gain phototransistor for compact, light-weight night vision intensifiers. The material system has kept improving in its crystal quality with low dislocation and defects with advanced material growth technique such as Molecular Beam Epitaxy. Their integration to silicon photoncis for optoelectronic devices has countless applications.

6.2 Prospects

As improved infrared sensitivity meets the novel computation scheme with image processing and neural network algorithms, the IR camera would extend its application to various edge devices monitoring health condition of human body, wildlife, and industrial safety. Moreover, the quest for semiconductor materials suitable for IR technology extends beyond traditional boundaries. Nanoscale engineering, nanogap plasmonic field enhancement, quantum dot integration, and metamaterial have pushed the limits of what is achievable in manipulating light at IR wavelengths, propelling the development of miniaturized, high-performance devices. As we stand at the precipice of technological innovation, the significance of semiconductor material engineering in IR technology cannot be overstated. It serves as the bedrock upon which innovative IR devices and applications flourish, offering solutions to contemporary challenges in healthcare, environmental monitoring, security, and communication. In the realm of optoelectronics, there is a growing interest in merging III-V materials with silicon, aiming to combine the strengths of both. This integration onto silicon substrates seeks to produce top-notch devices while capitalizing on the existing silicon manufacturing infrastructure. III-V semiconductors are pivotal in enabling high-frequency communication technologies, likely expanding their usage in photonics and high-speed communication tools, meeting the escalating demand for swift data transmission seen in advancements like 5G and beyond. Additionally, III-V materials play a central role in propelling quantum technologies such as computing, sensing, and communication. Their distinctive quantum properties position them as promising contenders for building essential components like qubits in quantum information processing. Concurrently, refining manufacturing techniques is crucial to scale up the production of superior-quality III-V materials at an economical level. Techniques like molecular beam epitaxy and metalorganic vapor phase epitaxy are undergoing enhancements to realize large-scale production [84, 85]. Integrating III-V materials with other semiconductor platforms presents challenges related to material compatibility, thermal expansion differences, and interface management. Ongoing research is dedicated to devising methods for seamless integration while preserving the distinctive attributes of III-V compounds. Moreover, cost remains a major obstacle hindering widespread adoption. Ongoing endeavors aim to devise cost-effective manufacturing processes and explore alternative substrates to enhance the economic viability of III-V technologies.

Beyond their traditional applications in electronics and optoelectronics, III-V compounds are branching into diverse fields like sensing technologies, biomedical devices, and even emerging domains like flexible electronics, owing to their unique properties and performance advantages. Ultimately, the future of III-V compound semiconductors looks promising, buoyed by continual research efforts focused on enhancing integration, refining manufacturing techniques, and exploring innovative applications across various technological domains. Looking ahead, the relentless pursuit of advancing semiconductor materials for IR technology promises a future brimming with breakthroughs. The convergence of materials science, nanotechnology, and photonics in engineering semiconductors for IR applications heralds a new era of transformative technologies, paving the way for enhanced quality of life, scientific discovery, and technological marvels yet to be imagined.

In essence, engineering semiconductor materials for IR technology epitomizes the synergy between scientific exploration and technological innovation, offering a glimpse into a future where the boundaries of what’s possible in sensing, imaging, and communication are continually redefined.

Bibliography

[1] M. L. Vatsia, K. Stich, and D. Dunlap, "Night-sky radiant sterance from 450 to 20 nanometers," *US Army Electronics Command*, vol. Fort Belvoir, Virginia, 1972.

[2] A. Rogalski, "History of infrared detectors," *Opto-electronics Review*, vol. 20, no. 3, pp. 279–308, 2012.

[3] K. S. NOVOSELOV, A. K. GEIMS, V. MOROZOV, D. JIANG, Y. ZHANG, S. V. DUBONOS, I. V. GRIGORIEVA, and A. A. FIRSOV, "Electric Field Effect in Atomically Thin Carbon Films," *Science*, vol. 306, no. 5696, pp. 666–669, 2004.

[4] H. Hertz, "Ueber den einfluss des ultravioletten lichtes auf die electrische entladung," *Annalen der Physik*, vol. 267, no. 8, pp. 983–1000, 1887.

[5] J. Elester and H. Geitel, "Ueber die entladung negativ electrischer korper durch das sonnen- und tageslicht," *Annalen der Physik*, pp. 497–514, 1889.

[6] F. Braun, "Uber die stromleitung durch schwefelmetalic," *Annalen der Physik and Chemie*, vol. 153, no. 4, pp. 556–563, 1874.

[7] J. Bose, "Detector for electrical disturbances," *US patent*, pp. 775,840, 1904.

[8] T. Case, "Notes on the change of resistance of certain substrates in light," *Physics Review*, vol. 9, pp. 305–310, 1917.

[9] G. Holst, J. de Boer, M. C. Teves, and C. F. Veenemans, “An apparatus for the transformation of light of long wavelength into light of short wavelength,” *Physica D: Nonlinear Phenomena*, vol. 1, pp. 297–305, 1934.

[10] S. ASAO, “Behaviour of the foreign metal particles in the composite photo-cathode,” *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, vol. 22, no. 6, pp. 448–486, 1940.

[11] S. Jonshon, “A history of light and colour measurement. science in the shadows,” *IOP Publishing Ltd, Bristol*, 2001.

[12] E. W. Kutzscher, “Review on detectors on infrared radiation,” *Electro-Optic System Design*, vol. 5, no. 30, 1973.

[13] P. Kruse, L. McGlauchlin, and R. McQuistan, “Elecments of infrared technolgy,” *Wiley, New York*, 1962.

[14] M. Judt and B. Ciesla, “Technology transfer out of germany after 1945,” *Routledge Studies in the History of Science, Technology and Medicine, Oversears Publishers Association, Amsterdam*, 1996.

[15] R. Cashman, “Film-type infrared photoconductors,” *Proceedings IRE*, vol. 47, pp. 1471–1475, 1959.

[16] D. Lovell, “Cashman thallous sulfide cell,” *Applied Optics*, vol. 10, pp. 1003–1008, 1971.

[17] D. Lovell, “The development of lead salt detectors,” *American Journal of Physics*, vol. 37, pp. 467–478, 1969.

[18] P. Norton, “Infrared detectors in the next millennium,” *Proceedings of SPIE*, vol. 3698, pp. 652–665, 1999.

[19] H. Tong, G. Arjavalingam, R. Haynes, G. Hyer, and J. Ritsko, “High-temperature thin-film pt–ir thermocouple with fast time response,” *Review of scientific instruments*, vol. 58, no. 5, pp. 875–877, 1987.

[20] K. Liddiard, “Thin-film resistance bolometer ir detectors,” *Infrared Physics*, vol. 24, no. 1, pp. 57–64, 1984.

[21] P. Richards, "Bolometers for infrared and millimeter waves," *Journal of Applied Physics*, vol. 76, no. 1, pp. 1–24, 1994.

[22] M. J. Golay, "Static multislit spectrometry and its application to the panoramic display of infrared spectra," *JOSA*, vol. 41, no. 7, pp. 468–472, 1951.

[23] K. Yamashita, A. Murata, and M. Okuyama, "Miniaturized infrared sensor using silicon diaphragm based on golay cell," *Sensors and Actuators A: Physical*, vol. 66, no. 1-3, pp. 29–32, 1998.

[24] R. Takayama, Y. Tomita, K. Iijima, and I. Ueda, "Preparation and characteristics of pyroelectric infrared sensors made of c-axis oriented la-modified pbti03 thin films," *Journal of applied physics*, vol. 61, no. 1, pp. 411–415, 1987.

[25] M. Moghavvemi and L. C. Seng, "Pyroelectric infrared sensor for intruder detection," in *2004 IEEE Region 10 Conference TENCON 2004.*, vol. 500, pp. 656–659, IEEE, 2004.

[26] C. Grimes, P. Richards, and S. Shapiro, "Josephson-effect far-infrared detector," *Journal of applied physics*, vol. 39, no. 8, pp. 3905–3912, 1968.

[27] E. D. Walsh, W. Jung, G.-H. Lee, D. K. Efetov, B.-I. Wu, K.-F. Huang, T. A. Ohki, T. Taniguchi, K. Watanabe, P. Kim, *et al.*, "Josephson junction infrared single-photon detector," *Science*, vol. 372, no. 6540, pp. 409–412, 2021.

[28] H. Elabd and W. F. Kosonocky, "Ptsi infrared schottky-barrier detectors with optical cavity," *Review*, vol. 43, p. 569, 1982.

[29] E. L. Stelzer, J. Schmit, and O. Tufte, "Mercury cadmium telluride as an infrared detector material," *IEEE Transactions on Electron Devices*, vol. 16, no. 10, pp. 880–884, 1969.

[30] L. Kidder, I. Levin, E. N. Lewis, V. Kleiman, and E. J. Heilweil, "Mercury cadmium telluride focal-plane array detection for mid-infrared fourier-transform spectroscopic imaging," *Optics letters*, vol. 22, no. 10, pp. 742–744, 1997.

[31] T. N. Krabach, C. O. Staller, S. N. Dejewski, T. J. Cunningham, M. Herring, and E. R. Fossum, "Ingaas detectors for miniature infrared instruments," in *Infrared and Millimeter-Wave Engineering*, vol. 1874, pp. 214–223, SPIE, 1993.

[32] A. Rogalski, "Infrared detectors, 2nd edition," *CRC Press, Boca Raton*, 2010.

[33] R. C. Jones, "Proposal of the detectivity d$**$ for detectors limited by radiation noise†," *J. Opt. Soc. Am.*, vol. 50, pp. 1058–1059, Nov 1960.

[34] M. Staninec, S. Douglas, C. Darling, K. Chan, H. Kang, R. Lee, and D. Fried, "Non-destructive clinical assessment of occlusal caries lesions using near-ir imaging methods," *Lasers in Surgery and Medicine*, vol. 43, no. 10, pp. 951–959, 2011.

[35] J. Simon, H. Kang, M. Staninec, A. Jang, K. Chan, C. Darling, R. Lee, and D. Fried, "Near-ir and cp-oct imaging of suspected occlusal caries lesions," *Lasers in Surgery and Medicine*, vol. 49, no. 3, pp. 215–224, 2017.

[36] B. Li, Q. Lin, and M. Li, "Frequency-angular resolving lidar using chip-scale acousto-optic beam steering," *Nature*, vol. 620, pp. 316–322, 2023.

[37] C. Wu, H. Yu, S. Lee, R. Peng, I. Takeuchi, and M. Li, "Programmable phase-change metasurfaces on waveguides for multimode photonic convolutional neural network," *Nature Communications*, vol. 12, no. 96, 2021.

[38] M. Halicek, G. Lu, J. V. Little, X. Wang, M. Patel, C. C. Griffith, M. W. El-Deiry, A. Y. Chen, and B. Fei, "Deep convolutional neural networks for classifying head and neck cancer using hyperspectral imaging," *Nano Letters*, vol. 17, no. 10, pp. 6315–6320, 2017.

[39] J. Bakker, L. Aleese, G. Meijer, and G. vonHelden, "Fingerprint ir spectroscopy to probe amino acid conformations in the gas phase," *Physical Review Letters*, vol. 91, no. 203003, 2003.

[40] C. Chen, S. Oh, and M. Li, "Coupled-mode theory for plasmonic resonators integrated with silicon waveguides towards mid-infrared spectroscopic sensing," *Optics Express*, vol. 28, no. 2, pp. 2020–2036, 2020.

[41] X. Ling, H. Wang, S. Huang, F. Xia, and M. Dresselhaus, "The renaissance of black phosphorus," *Perspective Engineering, PNAS*, vol. 112, no. 15, pp. 4524–4530, 2015.

[42] Z. Zhao, B. Wu, X. Wang, Z. Pan, Z. Liu, P. Zhang, X. Shen, Q. Nie, S. Dai, and R. Wang, "Mid-infrared supercontinuum covering 2.0–16 m in a low-loss telluride single-mode fiber," *Laser & Photonics Reviews*, vol. 11, no. 2, p. 1700005, 2017.

[43] K. F. Mak and J. Shan, "Photonics and optoelectronics of 2D semiconductor transition metal dichalcogenides," *Nature Photonics*, vol. 10, pp. 216–226, 2016.

[44] F. Xia, H. Wang, D. Xiao, M. Dubey, and A. Ramasubramaniam, "Two-dimensional material nanophotonics," *Nature Photonics*, vol. 8, pp. 899–907, 2014.

[45] R. Peng, A. Ripin, Y. Ye, J. Zhu, C. Wu, S. Lee, H. Li, T. Taniguchi, K. Watanabe, T. Cao, *et al.*, "Long-range transport of 2d excitons with acoustic waves," *Nature Communications*, vol. 13, no. 1, p. 1334, 2022.

[46] A. Ripin, R. Peng, X. Zhang, S. Chakravarthi, M. He, X. Xu, K.-M. Fu, T. Cao, and M. Li, "Tunable phononic coupling in excitonic quantum emitters," *Nature Nanotechnology*, pp. 1–7, 2023.

[47] Q. Zeng and Z. Liu, "Novel optoelectronic devices: transition-metal-dichalcogenide-based 2d heterostructures," *Advanced Electronic Materials*, vol. 4, no. 2, p. 1700335, 2018.

[48] A. K. Geim and K. S. Novoselov, "The rise of graphene," *Nature materials*, vol. 6, no. 3, pp. 183–191, 2007.

[49] A. K. Geim, "Graphene: status and prospects," *science*, vol. 324, no. 5934, pp. 1530–1534, 2009.

[50] J. M. Dawlaty, S. Shivaraman, J. Strait, P. George, M. Chandrashekhar, F. Rana, M. G. Spencer, D. Veksler, and Y. Chen, "Measurement of the optical absorption spectra of epitaxial graphene from terahertz to visible," *Applied Physics Letters*, vol. 93, no. 13, 2008.

[51] R. R. Nair, P. Blake, A. N. Grigorenko, K. S. Novoselov, T. J. Booth, T. Stauber, N. M. Peres, and A. K. Geim, "Fine structure constant defines visual transparency of graphene," *science*, vol. 320, no. 5881, pp. 1308–1308, 2008.

[52] A. B. Kuzmenko, E. Van Heumen, F. Carbone, and D. Van Der Marel, "Universal optical conductance of graphite," *Physical review letters*, vol. 100, no. 11, p. 117401, 2008.

[53] F. Bonaccorso, Z. Sun, T. Hasan, and A. Ferrari, "Graphene photonics and optoelectronics," *Nature photonics*, vol. 4, no. 9, pp. 611–622, 2010.

[54] X. Gan, R.-J. Shiue, Y. Gao, I. Meric, T. F. Heinz, K. Shepard, J. Hone, S. Assefa, and D. Englund, "Chip-integrated ultrafast graphene photodetector with high responsivity," *Nature photonics*, vol. 7, no. 11, pp. 883–887, 2013.

[55] K. Rana, J. Singh, and J.-H. Ahn, "A graphene-based transparent electrode for use in flexible optoelectronic devices," *Journal of Materials Chemistry C*, vol. 2, no. 15, pp. 2646–2656, 2014.

[56] J. Wu, H. A. Becerril, Z. Bao, Z. Liu, Y. Chen, and P. Peumans, "Organic solar cells with solution-processed graphene transparent electrodes," *Applied physics letters*, vol. 92, no. 26, 2008.

[57] X. Li, Y. Zhu, W. Cai, M. Borysiak, B. Han, D. Chen, R. D. Piner, L. Colombo, and R. S. Ruoff, "Transfer of large-area graphene films for high-performance transparent conductive electrodes," *Nano letters*, vol. 9, no. 12, pp. 4359–4363, 2009.

[58] B. Sensale-Rodriguez, R. Yan, M. M. Kelly, T. Fang, K. Tahy, W. S. Hwang, D. Jena, L. Liu, and H. G. Xing, "Broadband graphene terahertz modulators enabled by intraband transitions," *Nature communications*, vol. 3, no. 1, p. 780, 2012.

[59] J. Tong, M. Muthee, S.-Y. Chen, S. K. Yngvesson, and J. Yan, "Antenna enhanced graphene thz emitter and detector," *Nano Letters*, vol. 15, no. 8, pp. 5295–5301, 2015.

[60] N. Youngblood, C. Chen, S. J. Koester, and M. Li, "Waveguide-integrated black phosphorus photodetector with high responsivity and low dark current," *Nature Photonics*, vol. 9, pp. 247–252, 2015.

[61] N. Youngblood, Y. Anugrah, R. Ma, S. J. Koester, and M. Li, "Multifunctional graphene optical modulator and photodetector integrated on silicon waveguides," *Nano Letters*, vol. 14, no. 5, pp. 2741–2746, 2014. PMID: 24734877.

[62] L. Liu, C. Liu, L. Jiang, J. Li, Y. Ding, S. Wang, Y. G. Jiang, Y. Sung, J. Wang, S. Chen, D. Zhang, and P. Zhou, "Ultrafast non-volatile flash memory based on van der waals heterostructure," *Nature Nanotechnology*, vol. 16, pp. 874–881, 2021.

[63] L. Li, Y. Yu, G. J. Ye, Q. Ge, X. Ou, H. Wu, D. Feng, X. H. Chen, and Y. Zhang, "Black phosphorus field-effect transistors," *Nature Nanotechnology*, vol. 9, pp. 372–377, 2014.

[64] V. Tran, R. Soklaski, and Y. Liang, "Layer-controlled band gap and anisotropic excitons in few-layer black phosphorus," *Physical Review B*, vol. 89, no. 201408, 2014.

[65] G. ElMasry, D.-W. Sun, and P. Allen, "Near-infrared hyperspectral imaging for predicting colour, pH and tenderness of fresh beef," *Journal of Food Engineering*, vol. 110, no. 1, pp. 127–140, 2012.

[66] R. Peng, K. Khaliji, N. Youngblood, R. G. T. Low, and M. Li, "Midinfrared Electro-optic Modulation in Few-Layer Black Phosphorus," *Nano Letters*, vol. 17, no. 10, pp. 6315–6320, 2017.

[67] A. Gao, "Observation of ballistic avalanche phenomena in nanoscale vertical inse/bp heterostructures," *Nature Nanotechnology*, vol. 14, pp. 217–222, 2019.

[68] T.-Y. Chang, Y. Chen, D.-I. Luo, J.-X. Li, P.-L. Chen, S. Lee, Z. Fang, W.-Q. Li, Y.-Y. Zhang, M. Li, *et al.*, "Black phosphorus mid-infrared light-emitting diodes integrated with silicon photonic waveguides," *Nano Letters*, vol. 20, no. 9, pp. 6824–6830, 2020.

[69] J.-M. Liu, "Photonic devices," *Cambridge Press*, 2005.

[70] S. Nakamura, T. M. T. Mukai, and M. S. M. Senoh, "High-power gan pn junction blue-light-emitting diodes," *Japanese Journal of Applied Physics*, vol. 30, no. 12A, p. L1998, 1991.

[71] R. Breiter, W. A. Cabanski, K.-H. Mauk, W. Rode, and J. Ziegler, "Portable sequential multicolor thermal imager based on a mct 384 x 288 focal plane array," in *Infrared Technology and Applications XXVII*, vol. 4369, pp. 579–587, SPIE, 2001.

[72] N. Bar-Chaim, S. Margalit, A. Yariv, and I. Ury, "Gaas integrated optoelectronics," *IEEE Transactions on Electron Devices*, vol. 29, no. 9, pp. 1372–1381, 1982.

[73] C. Harder, B. V. Zeghbroeck, H. Meier, W. Patrick, and P. Vettiger, “5.2-ghz bandwidth monolithic gaas optoelectronic receiver,” *IEEE Electron Device Letters*, vol. 9, no. 4, pp. 171–173, 1988.

[74] C. Lee, “Picosecond optoelectronic switching in gaas,” *Applied Physics Letters*, vol. 30, no. 2, pp. 84–86, 1977.

[75] S. Chandrasekhar, “Optoelectronic system integration using inp-based hbts for lightwave communications,” *Solid-state Electronics*, vol. 41, no. 10, pp. 1413–1417, 1997.

[76] F. Leonberger and P. Moulton, “High-speed inp optoelectronic switch,” *Applied Physics Letters*, vol. 35, no. 9, pp. 712–714, 1979.

[77] A. Foyt, F. Leonberger, and R. Willamson, “Picosecond inp optoelectronic switches,” *Applied Physics Letters*, vol. 40, no. 6, pp. 447–449, 1982.

[78] A. Krost and A. Dadgar, “Gan-based optoelectronics on silicon substrates,” *Materials Science and Engineering:B*, vol. 93, no. 1, pp. 77–84, 2002.

[79] S. Denbaars, “Gallium-nitride-based materials for blue to ultraviolet optoelectronics devices,” *Proceedings of the IEEE*, vol. 85, no. 11, pp. 1740–1749, 1997.

[80] S. Keller and D. Steven P, “Gallium nitride based materials and their application for light emitting devices,” *Current Opinion in Solid State and Materials Science*, vol. 3, no. 1, pp. 45–50, 1998.

[81] Y. Arakawa, “Progress in gan-based quantum dots for optoelectronics applications,” *IEEE Journal of Selected Topic in Quantum Electronics*, vol. 8, no. 4, pp. 823–832, 2002.

[82] H. Mäntynen, N. Anttu, Z. Sun, and H. Lipsanen, “Single-photon sources with quantum dots in iii–v nanowires,” *Nanophotonics*, vol. 8, no. 5, pp. 747–769, 2019.

[83] P. Dhingra, P. Su, B. Li, R. Hool, A. Muhowski, M. Kim, D. Wasserman, J. Dallesasse, and M. Lee, “Low-threshold inp quantum dot and ingap quantum well visible lasers on silicon (001),” *Optica*, vol. 8, no. 11, pp. 1495–1500, 2021.

[84] P. Dhingra, A. J. Muhowski, B. D. Li, Y. Sun, R. D. Hool, D. Wasserman, and M. L. Lee, "Low-threshold visible InP quantum dot and InGaP quantum well lasers grown by molecular beam epitaxy," *Journal of Applied Physics*, vol. 133, p. 103101, 03 2023.

[85] B. Li, Y. Sun, R. D. Hool, and M. L. Lee, "Abrupt Te doping of GaInP grown by molecular beam epitaxy for solar cell applications," *Journal of Applied Physics*, vol. 134, p. 065702, 08 2023.

[86] R. Peng, "Electromechanically control the optoelectronic properties of 2d semiconductors," *Dissertation, University of Washington*, 2022.

[87] I. Coddington, N. Newbury, and W. Swann, "Dual-comb spectroscopy," *Optica*, vol. 3, no. 4, pp. 414–426, 2016.

[88] P. Lagage, "Exoplanets characterisation with the jwst and particularly miri," *European Planetary Science Congress*, pp. EPSC2015–757, 2015.

[89] R. Soref, "Mid-infrared photonics in silicon and germanium," *Nature Photonics*, vol. 4, pp. 495–497, 2010.

[90] J. Liang, Y. Hu, K. Zhang, Y. Wang, X. Song, A. Tao, Y. Liu, and Z. Jin, "2d layered black arsenic-phosphorus materials: Synthesis, properties, and device applications," *Nano Research*, pp. 1–16, 2022.

[91] M. Long, A. Gao, P. Wang, H. Xia, C. Ott, C. Pan, Y. Fu, E. Liu, X. Chen, W. Lu, T. Nilges, J. Xu, X. Wang, W. Hu, and F. Miao, "Room temperature high-detectivity mid-infrared photodetectors based on black arsenic phosphorus," *Science Advances*, vol. 3, no. 6, 2017.

[92] B. Seraphin and N. Bottka, "Franz-keldysh effect of the refracctive index in semiconductors," *Physical Review*, vol. 139, no. 2A, pp. 560–565, 1965.

[93] P. Chaisakul, D. Marris-Morini, M.-S. Rouifed, G. Isella, D. Chrastina, J. Frigerio, X. L. Roux, S. Edmond, J.-R. Coudevylle, and L. Vivien, "23 ghz ge/sige multiple quantum well electro-absorption modulator," *Opt. Express*, vol. 20, pp. 3219–3224, Jan 2012.

[94] T. Ido, S. Tanaka, M. Suzuki, M. Koizumi, H. Sano, and H. Inoue, “Ultra-high-speed multiple-quantum-well electro-absorption optical modulators with integrated waveguides,” *Journal of Lightwave Technology*, vol. 14, no. 9, pp. 2026–2034, 1996.

[95] T. Lian, K. Yang, X. Wang, M. Jiang, S. Sun, D. Niu, and D. Zhang, “Electro-absorption optical modulator based on graphene-buried polymer waveguides,” *IEEE Photonics Journal*, vol. 12, no. 4, pp. 1–10, 2020.

[96] J. Kim, S. S. Baik, S. H. Ryu, Y. Sohn, S. Park, B.-G. Park, J. Denlinger, Y. Yi, H. J. Choi, and K. S. Kim, “Observation of tunable band gap and anisotropic dirac semimetal state in black phosphorus,” *Science*, vol. 349, no. 6249, pp. 723–726, 2015.

[97] Z. Djuric, B. Livada, V. Jovic, M. Smiljanic, M. Matic, and Z. Lazic, “Quantum efficiency and responsivity of insb photodiodes utilizing the moss-burstein effect,” *Infrared Physics*, vol. 29, no. 1, pp. 1–7, 1989.

[98] J. Liu, M. Beals, A. Pomerene, S. Bernardis, R. Sun, J. Cheng, L. C. Kimerling, and J. Michel, “Waveguide-integrated, ultralow-energy gesi electro-absorption modulators,” *Nature photonics*, vol. 2, no. 7, pp. 433–437, 2008.

[99] D. Feng, S. Liao, H. Liang, J. Fong, B. Bijlani, R. Shafiiha, B. J. Luff, Y. Luo, J. Cunningham, A. V. Krishnamoorthy, *et al.*, “High speed gesi electro-absorption modulator at 1550 nm wavelength on soi waveguide,” *Optics express*, vol. 20, no. 20, pp. 22224–22232, 2012.

[100] X. Fu, J. Cheng, Q. Huang, Y. Hu, W. Xie, M. Tassaert, J. Verbist, K. Ma, J. Zhang, K. Chen, C. Zhang, Y. Shi, J. Bauwelinck, G. Roelkens, L. Liu, and S. He, “5 x 20 gb/s heterogeneously integrated iii-v on silicon electro-absorption modulator array with arrayed waveguide grating multiplexer,” *Opt. Express*, vol. 23, pp. 18686–18693, Jul 2015.

[101] M. Nedeljkovic, C. G. Littlejohns, A. Z. Khokhar, M. Banakar, W. Cao, J. S. Penades, D. T. Tran, F. Y. Gardes, D. J. Thomson, G. T. Reed, *et al.*, “Silicon-on-insulator free-carrier injection modulators for the mid-infrared,” *Optics Letters*, vol. 44, no. 4, pp. 915–918, 2019.

[102] H. Lin, Y. Song, Y. Huang, D. Kita, S. Deckoff-Jones, K. Wang, L. Li, J. Li, H. Zheng, Z. Luo, *et al.*, “Chalcogenide glass-on-graphene photonics,” *Nature Photonics*, vol. 11, no. 12, pp. 798–805, 2017.

[103] S. Lee, R. Peng, and M. Li, “Broadband, waveguide-integrated mid-infrared black phosphorus modulator with high modulation depth,” in *Conference on Lasers and Electro-Optics*, p. STh4L.2, Optica Publishing Group, 2020.

[104] M. Ono, M. Hata, M. Tsunekawa, K. Nozaki, H. Sumikura, H. Chiba, and M. Notomi, “Ultrafast and energy-efficient all-optical switching with graphene-loaded deep-subwavelength plasmonic waveguides,” *Nature Photonics*, vol. 14, no. 1, pp. 37–43, 2020.

[105] S. Lee, R. Peng, C. Wu, and M. Li, “Programmable black phosphorus image sensor for broadband optoelectronic edge computing,” *Nature communications*, vol. 13, no. 1, p. 1485, 2022.

[106] F. Rosi, C. Miliani, R. Braun, R. Harig, D. Sali, B. G. Brunetti, and A. Sgamellotti, “Noninvasive Analysis of Paintings by Mid-infrared Hyperspectral Imaging,” *Angewante Chemie*, vol. 52, no. 20, pp. 5258–5261, 2013.

[107] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, “Edge computing: Vision and challenges,” *IEEE Internet of Things Journal*, vol. 3, no. 5, pp. 637–646, 2016.

[108] M. Satyanarayanan, “The emergence of edge computing,” *Computer*, vol. 50, no. 1, pp. 30–39, 2017.

[109] D. Kim, C. Yu, S. Xie, Y. Chen, J.-Y. Kim, B. Kim, J. P. Kulkarni, and T. T.-H. Kim, “An overview of processing-in-memory circuits for artificial intelligence and machine learning,” *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 12, no. 2, pp. 338–353, 2022.

[110] H. Jang, C. Liu, H. Hinton, M.-H. Lee, H. Kim, M. Seol, H.-J. Shin, S. Park, and D. Ham, “An Atomically Thin Optoelectronic Machine Vision Processor ,” *Advanced Materials*, vol. 32, p. 2002431, 2020.

[111] L. Mennel, J. Symonowicz, S. Wachter, D. K. Polyushkin, A. J. Molina-Mendoza, and T. Mueller, “Ultrafast machine vision with 2D material neural network image sensors ,” *Advanced Materials*, vol. 32, p. 2002431, 2020.

[112] V. Gretsenko, T. V. Perevalov, and D. R. Islamov, “Electronic properties of hafnium oxide: A contribution from defects and traps,” *Physics Reports*, vol. 613, pp. 1–20, 2016.

[113] C. H. Chang and J. G. Hwu, “Trapping characteristics of al2o3/hfo2/sio2 stack structure prepared by low temperature in situ oxidation in dc sputtering,” *Journal of Applied Physics*, vol. 105, no. 094103, 2009.

[114] R. Tsu and L. Esaki, “Tunneling in a finite superlattice,” *Applied Physics Letter*, vol. 22, no. 562, 2003.

[115] T. Ahmed, “Fully light-controlled memory and neuromorphic computation in layered black phosphorus,” *Advanced Materials*, vol. 33, no. 2004207, 2021.

[116] F.Zhou and Y. Chai, “Near-sensor and in-sensor computing,” *Nature Electronics*, vol. 3, pp. 664–671, 2020.

[117] M. Nahmias, “Photonic multiply-accumulate operations for neural networks,” *IEEE J. Sel. Top. Quantum Electron., volume =*.

[118] F. Cai, “A fully integrated reprogrammable memristor–cmos system for efficient multiply–accumulate operations,” *Nature Electronics*, vol. 2, pp. 290–299, 2019.

[119] V. Camus, L. Mei, C. Enz, and M. Verhelst, “Review and benchmarking of precision-scalable multiply-accumulate unit architectures for embedded neural-network processing,” *IEEE J. Emerg. Sel.Top. Circuits Syst.*, vol. 9, pp. 697–711, 2019.

[120] B. Shen and I. K. Sethi, “Convolution-based edge detection for image/video in block dct domain,” *Journal of Visual Communication and Image Representation*, vol. 7, no. 4, pp. 411–423, 1996.

[121] N. Kanopoulos, N. Vasanthavada, and R. L. Baker, “Design of an image edge detection filter using the sobel operator,” *IEEE Journal of solid-state circuits*, vol. 23, no. 2, pp. 358–367, 1988.

[122] M. LeGallo, “Mixed-precision in-memory computing,” *Nature Electronics*, vol. 1, pp. 246–253, 2018.

[123] I. Giannopoulos, “8-bit precision in-memory multiplication with projected phase-change memory,” *IEEE Int. Electron Devices Meet. IEDM*, vol. 27, no. 7, pp. 1–4, 2018.

[124] A. Goda, “Recent progress on 3d nand flash technologies,” *Electron*, vol. 10, no. 3156, 2021.

[125] Y. Chai, “In-sensor computing for machine vision,” *Nature*, vol. 579, pp. 32–33, 2020.

[126] F. Liao, F. Zhou, and Y. Chai, “Neuromorphic vision sensors: Principle, progress and perspectives,” *J. Semidond.*, vol. 42, no. 013105, 2021.

[127] D. Ham, H. Park, S. Hwang, and K. Kim, “Neuromorphic electronics based on copying and pasting the brain,” *Nature Electronics*, vol. 4, pp. 635–644, 2021.

[128] J. Bullock, “Polarization-resolved black phosphorus/molybdenum disulfide mid-wave infrared photodiodes with high detectivity at room temperature,” *Nature Photonics*, vol. 12, pp. 601–607, 2018.

[129] H. Kim, “Actively variable-spectrum optoelectronics with black phosphorus,” *Nature*, vol. 596, pp. 232–237, 2021.

[130] G. Lu and B. Fei, “Medical hyperspectral imaging: A review,” *J. Biomed. Opt.*, vol. 19, no. 010901, 2014.

[131] M. Halicek, J. V. Little, X. Wang, A. Y. Chen, and B. Fei, “Optical biopsy of head and neck cancer using hyperspectral imaging and convolutional neural networks,” *J. Biomed. Opt.*, vol. 24, no. 036007, 2019.

[132] W. Yang, C. Yang, Z. Hao, C. Xie, and M. Li, “Diagnosis of plant cold damage based on hyperspectral imaging and convolutional neural network,” *IEEE Access*, vol. 7, pp. 118239–118248, 2019.

[133] W. Kong, C. Zhang, F. Liu, P. Nie, and Y. He, “Rice seed cultivar identification using near-infrared hyperspectral imaging and multivariate data analysis,” *Sensors*, vol. 13, pp. 8916–8927, 2013.

[134] G. Elmasry, D. W. Sun, and P. Allen, “Near-infrared hyperspectral imaging for predicting colour, ph and tenderness of fresh beef,” *J. Food Eng.*, vol. 110, pp. 127–140, 2012.

[135] L. Ravikanth, C. B. Singh, D. Jayas, and N. White, “Classification of contaminants from wheat using near-infrared hyperspectral imaging,” *Biosystem Engineering*, vol. 135, pp. 73–86, 2015.

[136] F. Rosi, "Noninvasive analysis of paintings by mid-infrared hyperspectral imaging," *Angewante Chemie Int. Ed*, vol. 52, pp. 5258–5261, 2013.

[137] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, pp. 844–853, 2017.

[138] S. Aryal, Z. Chen, and S. Tang, "Mobile hyperspectral imaging for material surface damage detection," *J. Comput. Civ. Eng.*, vol. 35, no. 04020057, 2020.

[139] Z. Wu, "Large-scale growth of few-layer two-dimensional black phosphorus," *Nature Materials*, vol. 20, pp. 1203–1209, 2021.

[140] S. Seo, S. Jo, S. Kim, J. Shim, S. Oh, J. Kim, K. Heo, J. Choi, C. Choi, S. Oh, D. Kuzum, H. Wong, and J. Park, "Artificial optic-neural synapse for colored and color-mixed pattern recognition," *Nature Communications*, vol. 9, no. 5106, 2018.

[141] Q. Feng, F. Yan, W. Luo, and K. Wang, "Charge trap memory based on few-layer black phosphorus," *Nanoscale*, vol. 8, pp. 2686–2692, 2016.

[142] D. Lee, Y. Choi, E. Hwang, M. Kang, S. Lee, and J. Cho, "Black phosphorus nonvolatile transistor memory," *Nanoscale*, vol. 8, pp. 9107–9112, 2016.

[143] L. Liu, C. Liu, L. Jiang, J. Li, Y. Ding, S. Wang, Y. Jiang, Y. Sun, J. Wang, S. Chen, D. Zhang, and P. Zhou, "Ultrafast non-volatile flash memory based on van der waals heterostructures," *Nature Nanotechnology*, vol. 16, pp. 874–881, 2021.

[144] L. Wu, A. Wang, J. Shi, J. Yan, Z. Zhou, C. Bian, and H. Gao, "Atomically sharp interface enabled ultrahigh-speed non-volatile memory devices," *Nature Nanotechnology*, vol. 16, pp. 882–887, 2021.

[145] H. Guo, L. Feng, and S. Xiang, "Research on an extended blue GaAs photocathode," in *Optoelectronic Devices and Integration* (H. Ming, X. Zhang, and M. Y. Chen, eds.), vol. 5644, pp. 579 – 582, International Society for Optics and Photonics, SPIE, 2005.

[146] J. Xu, L. Yan, Y. Cheng, K. Han, B. Liu, and T. Zhang, "Mechanism analysis of gen lll image intensifier gaas cathode photoelectric emission disability," in *International Symposium on Photoelectronic*

Detection and Imaging 2013: Low-Light-Level Technology and Applications, vol. 8912, pp. 241–247, SPIE, 2013.

[147] W. S. Enloe, R. Shelden, L. Reed, and A. Amith, "Electron-bombarded ccd image intensifier with a gaas photocathode," in *Electron Tubes and Image Intensifiers*, vol. 1655, pp. 41–49, SPIE, 1992.

[148] Z. Wang, Y. Zhang, S. Li, S. Li, J. Zhan, Y. Qian, F. Shi, H. Cheng, G. Jiao, and Y. Zeng, "Enhancement of near-infrared response for gaas-based photocathode with laminated graded-bandgap structure: theory and experiment," *Journal of Materials Research and Technology*, vol. 19, pp. 2008–2017, 2022.

[149] H. Yu, D. Kim, J. Lee, S. Baek, J. Lee, R. Singh, and F. So, "High-gain infrared-to-visible upconversion light-emitting phototransistors," *Nature Photonics*, vol. 10, pp. 129–134, 2016.

[150] J. Chen, D. Ban, M. Helander, Z. Lu, and P. Poole, "Near-infrared inorganic/organic optical upconverter with an external power efficiency of >100

[151] M. Gordon, "The lost history of trnasistor," *IEEE Spectrum*, pp. 44–49, 2004.

[152] H. Kroemer, "Theory of a wide-gap emitter for transistors," *Proceeding of IRE*, vol. 45, no. 11, pp. 1535–1537, 1957.

[153] J. C. Campbell, A. G. Dentai, C. A. Burrus, and J. F. Ferguson, "Inp-ingaas heterojunction phototransistors," *IEEE Journal of Quantum Electornics*, vol. QE-17, pp. 264–268, 1981.

[154] D. Fritzsche, E. Kuphal, and R. Aulbach, "Fast response inp/ingaasp heterojunction photoransistors," *Electronics Letter*, vol. 17, pp. 178–179, 1981.

[155] J. C. Campbell, C. A. Burrus, A. G. Dentai, and K. Ogawa, "Small-area high speed inp/ingaas phototransistor," *Applied Physics Letter*, vol. 39, pp. 820–821, 1981.

[156] A. Sasaki, K. J. Matsuda, Y. Kimura, and S. Fujita, "High-current ingaasp-inp phototransistors and some monolithic optical devices," *IEEE Trans. Electron Devices*, vol. ED-29, pp. 1382–1388, 1982.

[157] M. Kobayashi, S. Sakai, and M. Umeno, "Very low dark current p-p-p base ingaasp/inp phototransistors," *Japan Journal of Applied Physics*, vol. 22, pp. 159–161, 1983.

[158] Z. Hao, S. Han, F. Ren, B. Xiong, C. Sun, and Y. Luo, “Smooth and vertical etching of gaas/gainp/algainp using inductively coupled cl2/bcl3/ch4 plasma,” *Japanese Journal of Applied Physics*, vol. 43, no. 12, pp. 8304–8307, 2004.

[159] H. Ding, L. Lu, Z. Shi, D. Wang, L. Li, X. Li, Y. Ren, C. Liu, D. Cheng, H. Kim, N. C. Giebink, X. Wang, L. Yin, L. Zhao, M. Luo, and X. Sheng, “Microscale optoelectronic infrared-to-visible upconversion devices and their use as injectable light sources,” *Proceedings of the National Academy of Sciences*, vol. 115, no. 26, pp. 6632–6637, 2018.

[160] H. Luo, D. Ban, H. C. Liu, Z. R. Wasilewski, and M. Buchanan, “Optical upconverter with integrated heterojunction phototransistor and light-emitting diode,” *Applied Physics Letters*, vol. 88, p. 073501, 02 2006.

[161] M. A. Islam, P. Serles, B. Kumral, P. G. Demingos, T. Qureshi, A. Meiyazhagan, A. B. Puthirath, M. S. B. Abdullah, S. R. Faysal, P. M. Ajayan, D. Panesar, C. V. Singh, and T. Filleter, “Exfoliation mechanisms of 2D materials and their applications,” *Applied Physics Reviews*, vol. 9, p. 041301, 10 2022.

[162] Y. Huang, Y.-H. Pan, R. Yang, L.-H. Bao, L. Meng, H.-L. Luo, Y.-Q. Cai, G.-D. Liu, W.-J. Zhao, Z. Zhou, *et al.*, “Universal mechanical exfoliation of large-area 2d crystals,” *Nature communications*, vol. 11, no. 1, p. 2453, 2020.

[163] H. Liu, S. L. Wong, and D. Chi, “Cvd growth of mos2-based two-dimensional materials,” *Chemical Vapor Deposition*, vol. 21, no. 10-11-12, pp. 241–259, 2015.

[164] B. A. Naqvi, J. C. M. A. Shehzad, K. A. Min, M. F. Khan, S. Hussain, and J. Jung, “Visualizing degradation of black phosphorus using liquid crystals,” *Scientific Reports*, vol. 8, pp. 1–8, 2018.

Chapter A

Appendix One

A.1 Integration of 2D materials

The integration of 2D materials into silicon photonics or CMOS compatible platform requires the fabrication of large-scale, uniform, and highly crystalline 2D films, which is currently growing field of study. In general, two major methods of synthesizing and preparing 2D materials for optoelectronic devices consist of top-down methods and bottom-up methods. To be more specific, top-down methods mainly include micromechanical exfoliation and liquid phase exfoliation methods[[161, 162]], while bottom-up methods are comprised of chemical vapor deposition (CVD) and hydrothermal/solvothermal methods[[163]]. To take hBN and BP mechanical exfoliation for example, the hBN/BP heterostructure is picked up with a PC (polycarbonate) stamp and transferred on thermal oxide Silicon wafer. First, the hBN flake is transferred on SiO2 wafer, which is used for searching a target hBN flake and different from the final wafer, using scotch tape. BP flake is also prepared on another SiO2 wafer by scotch tape method. Then, each flake to be transferred is identified by optical microscope and AFM to make sure the thickness of the target flakes. After the flakes are confirmed, they are picked up in order by a PC stamp from the top layer to the bottom, in this case, hBN and then BP. The PC stamp consists of a film of PC, a PDMS pillar, and a glass slide. By slowly approaching this PC stamp while watching on a microscope, you can align the relative position of the flakes. The transfer of both flakes from their SiO2 wafer to the PC stamp is done at around 60 degree Celcius. Once the stack of hBN/BP is transferred on PC stamp, the whole structure is transferred onto the final Si wafer. After making a full contact of the stamp and the wafer, increase the temperature to 180 degree C to melt down the PC film. Finally, the PC is washed with Chloroform and IPA.

A.2 Fabrication Process for BP-PPT

Few-layer black phosphorus (bP) with 11 nm thickness was mechanically exfoliated from a bulk crystal (HQ Graphene) and dry transferred using a polydimethylsiloxane stamp onto a silicon substrate with 300 nm-thick thermally grown SiO2. The thickness of the bP flake was identified using an atomic force microscope (Bruker Dimensions Icon). The bP flake with the lateral size of 20 μm × 30 μm was patterned into an array of 4 × 3 pixels with each pixel in the size of 3 μm × 4 μm using electron beam lithography (EBL, JEOL- JBX6300FS) and inductively coupled plasma (ICP) etching based on SF6 chemistry when the ZEP 520 A resist was used as the protective mask. The source and drain contacts made of Ni/Au (5 nm/25 nm) were patterned by steps of EBL, electron beam evaporation, and lift-off in a solvent. The processes of bP exfoliation, thickness measurement, and lift-off of the deposited metal films were all performed in an Ar-filled glovebox with oxygen and water concentration <0.1 ppm to avoid the degradation of the bP flake due to the exposure to moisture and oxygen. Subsequently, the gate dielectric stack of Al2O3/HfO2/Al2O3 (AHA) was deposited by atomic layer deposition (ALD) systems. The 6 nm-thick tunneling layer of Al2O3 was grown on bP at 150 °C by thermal ALD, and the 7 nm-thick storage layer of HfO2 was then grown at 290 °C by plasma-enhanced ALD, followed by the 20 nm-thick blocking layer of Al2O3. The top gate electrode made of indium-tin-oxide (ITO was patterned by EBL and deposited by a pulsed sputtering system (Evatec LLS EVO).

A.2.1 Fabrication on the area with uniform thickness

Optical Microscope image of bP flake for A.2 (a) as transferred and (b) after patterning to the bP-PPT device array. All 12 devices are fabricated from the flat region of the bP flake as outlined with red-dashed line to minimize the device-to-device variation. Each device has source and drain contacts and an ITO top gate electrode. Inset: the thickness of bP by AFM along the white dashed line, showing 11 nm of the flake. Scale bar: 10 μm.

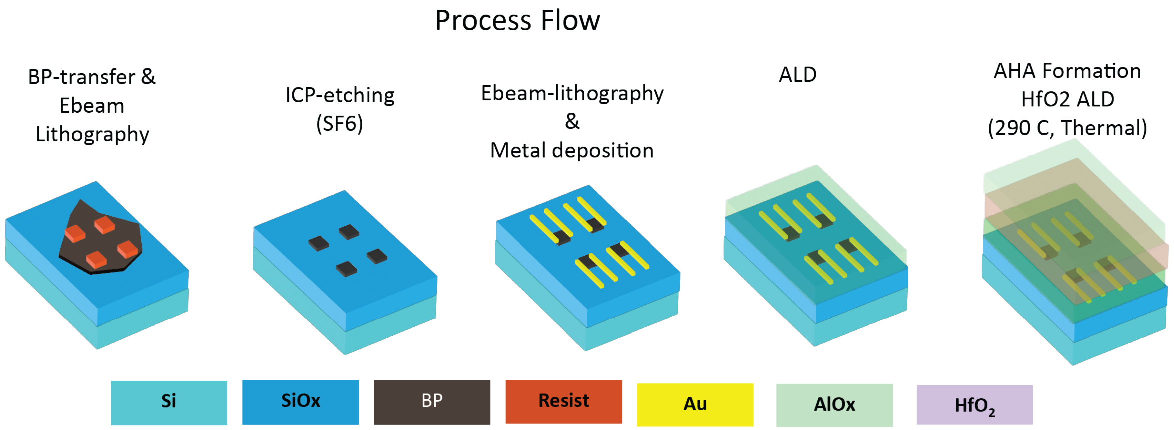


Figure A.1: 2D material integration and device fabrication.

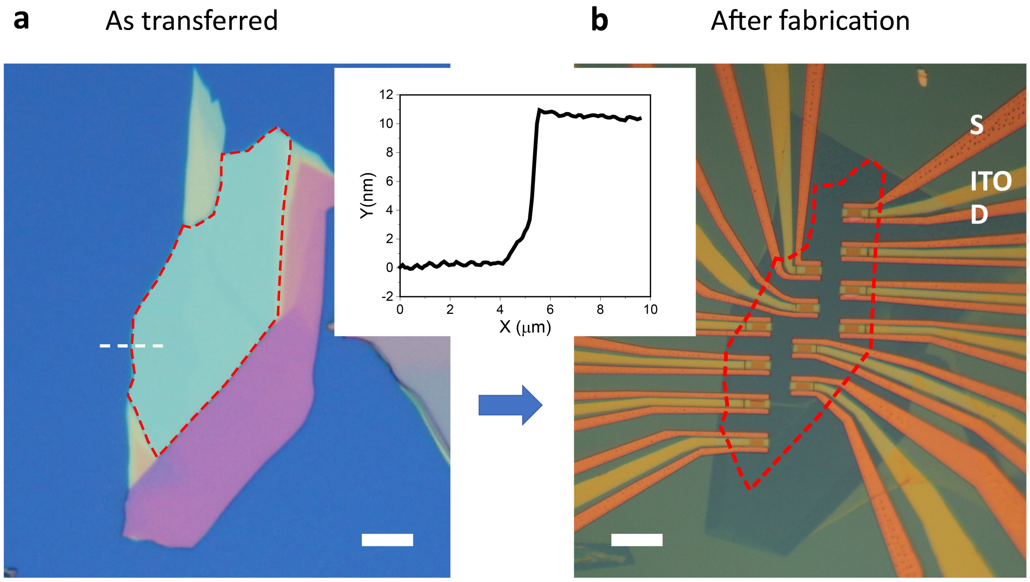


Figure A.2: 2D material integration and device fabrication. (a) Optical Microscope image of bP flake as transferred with red-dashed boundary indicating uniform thickness of 11 nm, which is measured by AFM along the white-dashed line. (b) The patterned bP-PPT device array. All 12 devices are fabricated from the flat region of the bP flake as outlined with red-dashed line to minimize the device-to-device variation. Each device has source and drain contacts and an ITO top gate electrode. Scale bar: 10 μm

A.2.2 Robustness of black phosphorus after Fabrication

Characterization of bP degradation during the fabrication process is performed by repeating the bP patterning process (e-beam lithography, ICP etch, resist removal). The optical microscopic images and AFM measurements of a bP flake as transferred and after etching are shown in Fig. A.3. The freshly exfoliated bP flake with a flat surface with 17-nm thickness is preserved in its optical contrast, surface smoothness, and thickness after the ICP etching. The thickness of the bP rectangles can be estimated considering the etched thickness of SiO_2 substrate as shown in the right column of Fig. A.3 a. Additionally, the Raman spectroscopy of the bP rectangle in Fig. A.3 b shows no signs of P_xO_y or H_xPO_y, at around Raman signal from 800 to 1000 cm^{-1}, revealing that oxidation and degradation of bP during the fabrication process is negligible[164].

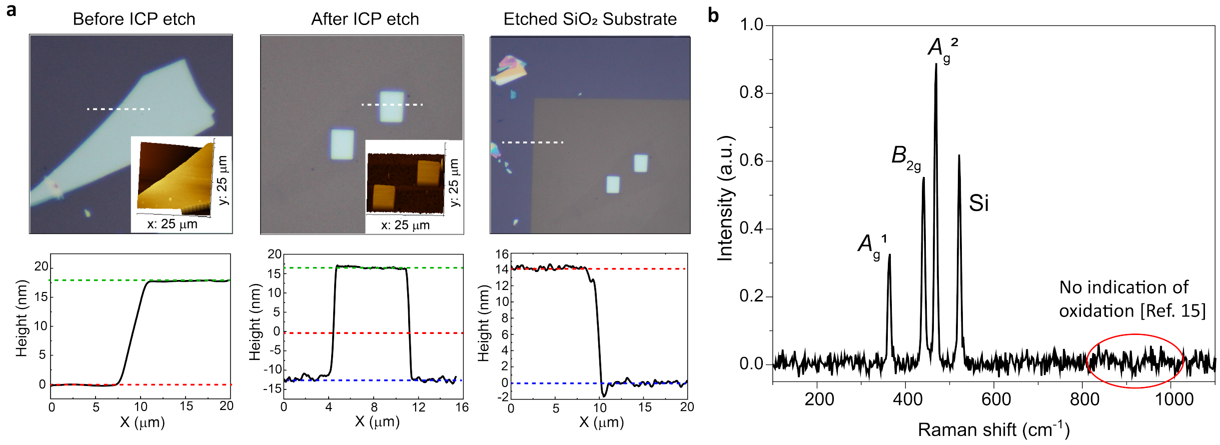


Figure A.3: (a) Optical images and AFM measurements for flat surface and preserved thickness of a bP flake upon ICP etching into two rectangles. The step height measurements are performed along the white dashed lines. The thickness of the bP flake is maintained as 17 nm. (b) Raman spectroscopy for ICP etched bP rectangle. It shows no indication of PxOy or HxPOy formed during the fabrication process [164].