Title
Two-dimensional materials as a new platform for atomically thin electronics and optoelectronics

Permalink
https://escholarship.org/uc/item/7r405223

Author
Cheng, Rui

Publication Date
2014

Peer reviewed|Thesis/dissertation
Two-dimensional materials as a new platform for atomically thin electronics and optoelectronics

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Material Science and Engineering

by

Rui Cheng

2014
ABSTRACT OF THE DISSERTATION
Two-dimensional materials as a new platform for atomically thin electronics and optoelectronics.

by

Rui Cheng

Doctor of Philosophy in Material Science and Engineering
University of California, Los Angeles, 2014
Professor Yu Huang, Chair

The discovery of graphene, made of single-layer carbon atoms, defines the starting point in the research and development of stable two-dimensional layer materials (2DLMs). Graphene has been one of the most extensively studied materials due to its unique band structure, the linear dispersion at the K point. It gives rise to novel phenomena, such as the anomalous quantum Hall effect, and has opened up a new category of “Fermi-Dirac” physics. Graphene has also attracted enormous attention for future electronics because of its exceptional high carrier mobility, high carrier saturation velocity, and large critical current density. Graphene's success has shown that other 2D materials beyond graphene may also exhibit fascinating properties and be used for accelerate the development of technology. Two-dimensional transition metal dichalcogenides (TMDs), such as molybdenum disulfide (MoS₂), are emerging as an exciting material system for future electronics due to their unique electronic properties and atomically thin geometry.

In this dissertation, I will firstly present my research studies on the first experimental observation of a dramatic enhancement of the conductance in a GNR field-effect transistor by a
perpendicular magnetic field. Very large negative MR of nearly 100% with conductance enhanced over 10,000 times was observed at low temperatures; and more than 50% remained at room temperature. Then I will show a new approach for the scalable fabrication of high performance graphene transistors with transferred gate stacks. This unique device structure enables scalable fabrication of self-aligned graphene transistors with unprecedented performance including a record high cut-off frequency up to 427 GHz. Taking a step further, I will demonstrate the best performed MoS$_2$ transistors with an on-off ratio exceeding $10^7$, excellent current saturation and a highest intrinsic gain over 30. On-chip microwave measurements demonstrate a highest intrinsic cut-off frequency $f_T$ of 42 GHz and a maximum oscillation frequency $f_{\text{MAX}}$ of 50 GHz. Furthermore, I will, for the first time, present the integration of multiple MoS$_2$ transistors on quartz and flexible substrates to form a logic inverter or radio frequency amplifier with voltage gain in the gigahertz regime. Finally I will present the first demonstration of an atomically thin heterojunction p-n diode by vertically stacking p-type monolayer tungsten diselenide (WSe$_2$) and n-type few-layer molybdenum disulfide (MoS$_2$). Electrical measurement demonstrates excellent diode characteristics with well defined current rectification behaviour, and photocurrent mapping shows clear photoresponse in the entire overlapping region with an external quantum efficiency as large as 12 %. Electroluminescence studies show prominent emission with both excitonic and hot electron luminescence peaks. A systematic investigation of the emission spectra reveals distinct layer-number dependent and temperature dependent characteristics and can offer important insight about the electron-orbital interaction in the layered materials. These findings can open up exciting new opportunities for 2DLMs in the application of electronics and optoelectronics.
The dissertation of Rui Cheng is approved.

Yong Cheng
Yang Yang
Xiangfeng Duan

Yu Huang, Committee Chair

University of California, Los Angeles

2014
List of Figures

Figure 1-1 | Distinguish single layer graphene .................................................................11
Figure 1-2 | Graphene band structure ..............................................................................12
Figure 2-1 | Experimental measured conduction band gaps $E_g$ as a function
of ribbon widths with different sets of crystal orientation ..................................33
Figure 2-2 | Anderson localization induced conductance suppression in edge
disordered AGNRs ................................................................................................34
Figure 2-3 | Schematic of quantum dots formation in GNRs due to disordered
potentials and confinement gaps .........................................................................35
Figure 2-4 | Experimental demonstration of Coulomb blockade diamond like
transport features in GNR FETs .........................................................................36
Figure 2-5 | Contour graph for localized edge spin state in zigzag GNR .......................37
Figure 2-6 | Electrical transport measurement of a graphene nanoribbon-FET ...............38
Figure 2-7 | Tunable magnetoresistance in graphene nanoribbon-FET ..........................39
Figure 2-8 | Current ratio $I(8T)/I(0T)$ as a function of source-drain voltage ..............40
Figure 2-9 | Temperature dependent magneto-transport properties .........................41
Figure 2-10 | Magneto-transport properties of a short channel
graphene nanoribbon-FET .............................................................................42
Figure 2-11 | Evolution of source-drain gap at selected gate voltages for
the short channel graphene nanoribbon FET ..................................................43
Figure 2-12 | Current-voltage characteristics under in-plane magnetic field ...............44
Figure 2-13 | Temperature dependence of the minimum conductance
at different magnetic field ................................................................................45
Figure 3-1 | Schematic illustration of the fabrication of self-aligned
Figure 3-2| The self-aligned graphene transistor .................................................................56

Figure 3-3| Room-temperature electrical characteristics of the CVD grown graphene transistors .................................................................57

Figure 3-4| RF performance of self-aligned CVD graphene transistors ..................................................58

Figure 3-5| RF performance of self-aligned peeled graphene transistors .................................................................59

Figure 3-6| The Raman spectrum of CVD grown graphene on SiO₂/Si substrate .................................................................59

Figure 3-7| Gate-leakage current versus top-gate voltage .................................................................59

Figure 3-8| The back and forth sweep of Iₜₒ₋Vₜ₉ curve of a peeled graphene device with transferred gate stack .................................................................60

Figure 3-9| Finite element simulation of the top-gate capacitance .................................................................60

Figure 3-10| Plot of power gain versus frequency .................................................................60

Figure 4-1| Schematic illustration and characterization of the MoS₂ FETs with transferred gate stacks .................................................................78

Figure 4-2| DC characterization of the self-aligned MoS₂ FETs .................................................................79

Figure 4-3| Radio frequency performance of the self-aligned MoS₂ FETs .................................................................80

Figure 4-4| Demonstration of integrated logic inverter and signal amplifier based on few-layer MoS₂ FETs on quartz substrate .................................................................81

Figure 4-4| Demonstration of integrated logic inverter and signal amplifier based on few-layer MoS₂ FETs on flexible substrate .................................................................82

Figure 4-6| Optical microscopy, atomic force microscopy, and Raman microscopy characterization of MoS₂ on 300 nm SiO₂ substrate .................................................................83

Figure 4-7| The Iₜₒ₋Vₜ₉ transfer characteristic curves for back-gated MoS₂ transistors with different thickness .................................................................84
Figure 4-8| Finite element simulation of electrostatic capacitance between the gate stack and MoS$_2$ in a 68 nm top-gate device………………………………………..85

Figure 5-1| Schematic illustration and band diagram of the WSe$_2$/MoS$_2$ vertical heterojunction p-n diode……………………………………………………………..101

Figure 5-2| Structural characterization of WSe$_2$/MoS$_2$ heterojunction p-n diode……………………………………………………………..102

Figure 5-3| Electrical characterization of WSe$_2$/MoS$_2$ heterojunction p-n diode……………………………………………………………..103

Figure 5-4| Photoresponse of the WSe$_2$/MoS$_2$ heterojunction p-n diode……………………………………………………………..104

Figure 5-5| Electroluminescence from WSe$_2$/MoS$_2$ heterojunction p-n diode……………………………………………………………..105

Figure 5-6| The AFM and Raman characterizations of synthetic WSe$_2$ domain……………………………………………………………..106

Figure 5-7| The analysis and peak fittings for the EL spectra of both ML- and BL-WSe$_2$/MoS$_2$ heterojunction……………………………………………………………..106
Acknowledgements

Foremost, I would like to express my sincere gratitude to my advisor Prof. Yu Huang and Prof. Xiangfeng Duan for their patience, motivation, enthusiasm, and immense knowledge. Your guidance helped me in all the time of my research and writing of this thesis. I could not have imagined having a better advisor and mentor for my Ph.D study. Besides my advisors, I would like to thank the rest of my thesis committee: Prof. Yang Yang and Prof. Yong Chen, for their encouragement, talented comments, and insightful questions. Special thanks go to Tom, Minji, Yuwei, Hoc and other staffs in UCLA NRF group for the support in clean room facilities.

I am also grateful to my labmates from UCLA MSE department and Chemistry department: Jingwei, Lei, Dehui, Yuan, Chen, Yungchen, Hailong, Lixin, Yongquan, Yu, Hao, Jonathan, Ben, Di, Yujing, Nathan, Gang, Kayla, Chin-Yi, Hua, Teng, Linyan, Yongjia, Xing, Qiao, Ivan, Tahani, Enbo, Sen, Yang and Woojong, for the experimental support and enlightening suggestions. I would like to thank my UCLA friend Mingshen, Caifu, Jianshi, Guangyu, Murong, Letian, Yue, Yanjie and Ming. All of you make my Ph.D life much more delightful!

Last but not least, I would like to thank my family, my parents, my parents in law and my wonderful wife, Shan. Thank you for the continuous support you have given me throughout my time in graduate school. I always know that you believed in me and want the best for me. You selflessly encouraged me to explore new directions in life and seek my own destiny. This journey would not have been possible if not for you, and I dedicate this milestone to you.
Biography

Rui Cheng received his B.S. degree in Physics from Peking University, Beijing, China in 2009. He joined Dr. Yu Huang’s group in 2009 and conducted the research of two-dimensional materials as a new platform for atomically thin electronics, opto-electronics. He received his M.S. degree in Materials Science and Engineering, University of California, Los Angeles in 2010. He was awarded UCLA graduate Fellowship in 2009. He has 6 first-author and 22 co-author papers published in peer-review journals. He gave oral presentations in Material Research Society Meeting 2013 Fall in Boston. His research interests include electrical and optical properties of two-dimensional layered materials or other nanomaterials.
INTRODUCTION

Graphene is two-dimensional crystalline form of carbon: a single layer of carbon atoms arranged in hexagons, like a honeycomb. This 2D structure comes from triangular planar arrangement of carbon-carbon $\sigma$ bonds due to $sp^2$ hybridization, which is further enforced by half-filled $\pi$ band\(^1\). This strong carbon-carbon bonding enables the lattice stability even up to 1500 °C upon annealing\(^2-^3\). Although graphene may be easily produced with pencil writing, it is difficult to locate such atomic layer thin material by conventional imaging techniques such as atomic force microscopy (AFM), transmission electron microscopy (TEM) and scanning electron microscopy (SEM), due to the low through-put and uncertainty of sample preparation\(^4-^5\). This came to an end in 2004 when Geim and co-workers successfully isolated atomic thin carbon layer by micro-mechanical cleavage method\(^4\). The discovery of graphene owed to its unique optical property: this atomic thin material can be visualized using optical microscope due to its notable opacity when prepared on Si substrate with certain thickness of SiO\(_2\) (Fig. 1-1a). Later on, it was discovered that graphene displayed distinctive Raman spectral characteristics, which was very powerful in distinguishing between single, double and multi-layer graphene (Fig. 1-1b)\(^9-^12\). The simple technique of optical microscopy in conjunction with Raman spectroscopy greatly facilitated the scientific research on graphene based novel materials.

Ever since it was first experimentally discovered, graphene is quickly rising to be one of the most attractive material systems for fundamental studies as well as potential applications due to its unique electronic properties. The band structure of single layer graphene is determined to have 2D character and a linear dispersion relation of electronic wave functions with perfect electron-hole symmetry, in which Fermi surface consists of two cones touching at one singular, so-called Dirac point, where the density of states is zero (Fig. 1-2a inset)\(^1\). The first demonstrated graphene
device showed ambipolar field effect - the resistance showed a peak as a function of gate voltage where the charge carrier changed the sign at the conductance minimum, beside which the conductance increased linearly with gate voltage on both sides of the resistance peak (Fig. 1-2a)\textsuperscript{5,13}.

The in depth study of Shubnikov-de Haas oscillation revives the nature that charge carriers in single layer graphene is massless Dirac fermions- similar to relativistic particles with zero rest mass with a Fermi velocity about $10^{-6} \text{m/s}$\textsuperscript{14}. The observed charge carrier mobility can reach 200,000 cm$^2$V$^{-1}$s$^{-1}$ for both electrons and holes even at room temperature\textsuperscript{15}. Unlike some other high mobility semiconductors, such as InSb which can only exhibit high mobility (mobility exceeding 60,000 cm$^2$V$^{-1}$s$^{-1}$ at room temperature) at a very low doping level with a 2D electron gas carrier concentration of $10^{11}$ cm$^{-2}$\textsuperscript{16,17}, graphene retains high mobility even with electron or hole doping with carrier density $>10^{12}$ cm$^{-2}$\textsuperscript{15,18-20}, enabling high current delivery through the device. The observation of quantum hall effect at room temperature also indicated exceptional electronic quality of graphene crystals\textsuperscript{14,21-23}. These extraordinary electrical properties make graphene a good candidate for high frequency device applications. Wafer scale gigahertz graphene transistors with cut-off frequency as high as 100 GHz were demonstrated\textsuperscript{24,25}. This record recently increased to 212 GHz in CVD grown graphene and to 427 GHz in mechanical exfoliated graphene by using a transfer gate technique\textsuperscript{26,27}. These studies define a new pathway to large scale fabrication of high performance graphene transistors, and holds significant potential for future application of graphene based devices in ultra-high frequency circuits.

Graphene has been explored for ultrahigh speed transistors with the intrinsic cut-off frequency exceeding 400 GHz, but typically with insufficient current on-off ratio and little voltage gain due to its zero band gap semimetal nature. Alternatively, other two-dimensional layered materials (2DLMs), such as molybdenum disulfide (MoS$\textsubscript{2}$), are emerging as an exciting material
system for future electronics due to their unique electronic properties and atomically thin geometry\textsuperscript{28-35}. The semiconducting MoS\textsubscript{2} have recently drawn considerable interest for overcome these disadvantages of graphene to enable atomically thin transistors with high on-off ratio and intrinsic voltage gain\textsuperscript{36-43}. A single layer of MoS\textsubscript{2} consists of a layer of Mo atoms sandwiched between two layers of S atoms. As a 2D material, it shares many interesting characteristics of the well-known graphene such as atomically thin thickness, excellent electronic properties, high mechanical flexibility, and partial optical transparency\textsuperscript{28-30}. With a direct band gap of 1.8 eV \textsuperscript{34,35}, it overcomes the key shortcomings of graphene for electronic applications — the lack of band gap and obvious current saturation\textsuperscript{44}. 2D electronics based on single or few-layer MoS\textsubscript{2} represents the ultimate limit of thickness for pushing the limits of the Moore’s Law. With a larger band gap than silicon, MoS\textsubscript{2} is also advantageous for suppressing the source-to-drain tunnelling current in transistors at the scaling limit\textsuperscript{45}. Moreover, MoS\textsubscript{2} and other transition metal dichalcogenide (TMD) materials are attractive as an alternative material for low-cost flexible electronics that is currently dominated by amorphous silicon and organic semiconductors with low mobilities around 1 cm\textsuperscript{2}/V\textperiodcentered s or less\textsuperscript{46-48}.

Two-dimensional layered materials, such as graphene, MoS\textsubscript{2}, and WSe\textsubscript{2}, are emerging as an exciting material system for a new generation of atomically thin optoelectronics, including photodetectors\textsuperscript{49-55}, ultrafast lasers\textsuperscript{56}, polarizers\textsuperscript{57}, touch panels\textsuperscript{58} and optical modulators\textsuperscript{59} due to their unique electronic and optical properties\textsuperscript{60-71}. In this regard, the monolayer transition metal dichalcogenides (ML-TMDs) is particularly interesting due to their direct energy bandgap and the non-centrosymmetric lattice structure\textsuperscript{60,61}. The p-n diodes represent the most fundamental device building block for most optoelectronic functions, including photodiodes and light emitting diodes. However, it is particularly difficult to create p-n diodes in atomically thin TMDs due to the
inability to selectively dope them into p- or n-type semiconductors. Electroluminescence (EL) from ML-MoS$_2$ has been reported in a metal-MoS$_2$ Schottky junction through a hot carrier process. Electrostatic doping has also been used to create planar p-n diodes, but usually with relatively gradual doping profile (limited by the fringe electrical field) and typically relatively low optoelectronic efficiency (e.g. photon to electron conversion external quantum efficiency (EQE) $\sim 0.1-1\%$). The atomically thin geometry of these 2D materials can allow for band structure modulation in a vertical direction as well as forming atomically sharp heterojunctions. For example, this strategy allows gapless graphene to be used in field-effect tunneling devices, barristers, inverters, and photodetectors while stacked together with other 2D materials in the vertical direction. Although the nearly perfect 2D structure and low density of states in graphene provide advantages in some heterostructure devices, its gapless nature prevents the formation of a large potential barrier for charge separation and current rectification. The vertical heterojunction p-n diode formed between one TMD material and a bulk material has recently been reported, but usually with no EL or very weak EL.

In this thesis, I will first present my research studies on the first experimental observation of a dramatic enhancement of the conductance in a GNR field-effect transistor by a perpendicular magnetic field. Very large negative MR of nearly 100% with conductance enhanced over 10,000 times was observed at low temperatures; and more than 50% remained at room temperature. Then I will show a new approach for the scalable fabrication of high performance graphene transistors with transferred gate stacks. This unique device structure enables scalable fabrication of self-aligned graphene transistors with unprecedented performance including a record high cut-off frequency up to 427 GHz. Taking a step further, I will demonstrate the best performed MoS$_2$ transistors with an on-off ratio exceeding $10^7$, excellent current saturation and a highest intrinsic
gain over 30. On-chip microwave measurements demonstrate a highest intrinsic cut-off frequency $f_T$ of 42 GHz and a maximum oscillation frequency $f_{MAX}$ of 50 GHz. Furthermore, I will, for the first time, present the integration of multiple MoS$_2$ transistors on quartz and flexible substrates to form a logic inverter or radio frequency amplifier with voltage gain in the gigahertz regime. Finally I will present the first demonstration of an atomically thin heterojunction p-n diode by vertically stacking p-type monolayer tungsten diselenide (WSe$_2$) and n-type few-layer molybdenum disulfide (MoS$_2$). Electrical measurement demonstrates excellent diode characteristics with well defined current rectification behaviour, and photocurrent mapping shows clear photoresponse in the entire overlapping region with an external quantum efficiency as large as 12 %. Electroluminescence studies show prominent emission with both excitonic and hot electron luminescence peaks. A systematic investigation of the emission spectra reveals distinct layer-number dependent and temperature dependent characteristics and can offer important insight about the electron-orbital interaction in the layered materials. These findings can open up exciting new opportunities for 2DLMs in the application of electronics and optoelectronics.
References


19. Lemme, M. C.; Echtermeyer, T. J.; Baus, M., and Kurz, H., Non-volatile switching in


37. Chang, H. Y. *et al.* High-performance, highly bendable MoS2 transistors with high-k


77. Georgiou, T. et al. Vertical field-effect transistor based on graphene-WS2 heterostructures for


Figures & Legends

**Figure 1-1| Distinguish single layer graphene.** a, Optical microscopy of mechanical peeled graphene on top of Si wafer with 300 nm SiO$_2$ (Adapt from [14]). b, Raman spectra of single layer graphene compare to bulk graphite at 514 nm laser excitation. The two most intensive features for graphite and graphene are the G peak at 1580 cm$^{-1}$ and a band $\sim$2700 cm$^{-1}$ called 2D peak. The 2D peak of graphite consist 2 components (showing as peak with a shoulder) which are roughly 1/4 and 1/2 height of the G peak; As a comparison, the graphene 2D peak is a single and sharp peak roughly 4 times more intense than G peak. (Adapted from [9]).
Figure 1-2 | Graphene band structure. a, Schematic of graphene band structure and ambipolar electric field effect in single layer graphene. b, Statistics of universal non-zero minimal conductivity at zero average carrier density. The solid dots show an improvement of device homogeneity with decrease of minimum conductivity approaching a value of $4e^2/h$. (Adapted from [5])
Chapter I: VERY LARGE MAGNETORESISTANCE IN GRAPHENE NANORIBBONS

A. Introduction to electrical transport in non-ideal GNRs

It was shown that most experimentally fabricated GNRs FETs were p-type transistors with charge neutrality point moved to the positive gate region, and with either slightly or heavily suppressed electron transport\textsuperscript{1}. This is primarily due to the absorption of charge impurities and oxidation of carbon-carbon bond during the fabrication process. Similar to bulk graphene, annealing GNR devices in vacuum and H\textsubscript{2} gas environment or removing polymer resist etching mask was proved to be effective to reduce impurity doping so that ambipolar transistor behavior can be partially recovered\textsuperscript{2}. Avouris and coworkers have fabricated GNR FETs with strong electron transport branch and negative shift of charge neutrality point by potassium doping and annealing\textsuperscript{3}. These studies indicated that the performance of a GNR-FET was heavily dependent on doping and impurities. At the same time, the room temperature on-off ratio of the devices all scaled inversely with the ribbon width: the FETs with width of 10-15nm typically had an on-off ratio around 10, and increased sharply when the width shrinks down to sub-10nm regime\textsuperscript{4,5}. The highest room temperature on-off ratio of $10^7$ for a GNR-FET was achieved in sonochemical synthesized GNRs with ribbon width in the sub-5nm regime\textsuperscript{4,6}.

The reported device mobility varied from one research group to another depending on different fabrication strategies. In general, the room temperature charge transport in GNR device is limited by intrinsic scattering, phonon scattering from supporting substrate, impurity scattering and line edge roughness scattering\textsuperscript{7-10}. Carrier mobility of $2,700 \text{ cm}^2\text{V}^{-1}\text{s}^{-1}$ was reported when ribbon width was larger than 56 nm in e-beam lithography fabricated GNR FETs\textsuperscript{11}. This value is close to that reported in top gated bulk graphene device without sophisticated treatment\textsuperscript{12,13}, indicating that room temperature charge transport in wide GNRs is limited by impurities scattering. The carrier
mobility starts to decrease when the ribbon width shrinks as an effect of the line edge roughness scattering\textsuperscript{8,10}. For e-beam lithography fabricated GNR FETs, the mobility decreased to less than 200 cm\textsuperscript{2}V\textsuperscript{-1}s\textsuperscript{-1} when the ribbon width was less than 20 nm, and a linear extrapolation gave a mobility value of < 10 cm\textsuperscript{2}V\textsuperscript{-1}s\textsuperscript{-1} for 2 nm GNRs. This is significantly smaller than the observed mobilities of chemically derived GNRs (~200 cm\textsuperscript{2}V\textsuperscript{-1}s\textsuperscript{-1} with a ribbon width of ~2 nm)\textsuperscript{6}, and those observed in the etched GNRs using nanowire mask (~ 1300 cm\textsuperscript{2}V\textsuperscript{-1}s\textsuperscript{-1} with a ribbon width of ~15 nm\textsuperscript{14}, indicating much smoother edges and less impurity scattering in the latter cases.

Kim and co-workers studied the low temperature electrical transport of GNRs FETs and showed diamond of suppressing conductance at low temperature of 1.6K\textsuperscript{15,16}, similar to other non-zero band gap semiconductor materials. The equivalent band gap scaled inversely proportional to ribbon width with a fitting of $\alpha/(W-W^*)$ (Figure 2-1) where $\alpha$ was scaling constant with a value about 0.2 meVnm and $W^*$ corresponds to an inactive width. The scaling law was similar to the theoretical calculation for ideal armchair GNRs with $\alpha$ in the range of 0.2-1.5 meVnm. However, their studies showed no crystallography dependence of the measured equivalent band gap, as shown in Figure 2-1 inset\textsuperscript{15}, which is quite different from theoretical predictions for ideal ribbons. Furthermore, low temperature measurement observed conduction resonance peaks inside the transport suppressed gaps, which was difficult to explain using a general band gap theory. These divergences suggested that mechanisms other than direct band gap may be accounted for the semiconducting behavior in experimental fabricated GNR devices.

Until now, the influence of edge structure on the electronic property can only be verified with STS study of local density of state\textsuperscript{17,18}. The STS studies of GNRs and quantum dots showed the existence of energy gaps for predominantly armchair-edge nanostructures. The measured energy gaps became smaller with increasing fraction of the zigzag edge, and became metallic when
zigzag edge dominated\textsuperscript{18}. These results agreed well with theoretical calculations for ideal GNRs, although the STS and STM is limited for real device applications.

**B. Edge roughness induced localization effect**

Motivated by the inconsistence between the early theoretical predications of ideal GNRs and the recently observed experimental phenomena, theoretical calculations based on non-ideal GNRs were reported in which edge roughness was taken into consideration as a mimic of real device morphology (Figure 2-2). The simulations which employed tight-binding model showed that the transfer of electron wave function can be largely hindered due to edge disorder induced quasi-one-dimensional Anderson localization so that the suppression of conductance near the charge neutrality point would occur even with moderate edge roughness\textsuperscript{19,20}. The calculation of the local density of state showed that the edge disorder can induce strong enhancement of the electron density at the edge to form surface like localized states which do not participate in the transport\textsuperscript{21,22}. The extracted localization length can vary with the disorder type, and decrease with the increase of disorder amplitude and with the shrinkage of ribbon width\textsuperscript{19, 22-24}. Also, the localization length reaches its minimum when the energy level approaches Dirac point, resulting in enlarged transport gaps\textsuperscript{20,24}. As shown in Figure 2-2b and c, a transport gap would develop where all the conducting channels were suppressed with the increasing disorder probability and ribbon length. In the tight-binding model, the Anderson localization induced transport band gap was found not only in the edge disordered armchair GNRs, but also in zigzag GNRs\textsuperscript{21,24}. The suppression of orientation effect was also predicted with intensive disorder in which two rows of edge atoms were involved\textsuperscript{25}. The calculated transport gaps also scaled with ribbon width for a GNR with moderate edge roughness with a fitting of \( E_{tg} = AW^{\alpha} \), where \( A \) was a constant, \( W \) was the ribbon width and \( \alpha \approx -1 \), which was in the similar form to the experimental report in the case of e-beam lithography.
fabricated GNRs. For FET application, the edge disorder will significantly reduce on-state current due to back scattering of electron waves, and increase off-state current due to the increasing band to band tunneling\textsuperscript{26}. The device performance may also vary from sample to sample if the detailed edge defects configuration is not controlled\textsuperscript{26}. These results point out potential challenges for the production of uniform GNR devices.

C. Coulomb blockade effect

Another important factor in disordered GNRs has been considered was the Coulomb blockade effect. The Coulomb blockade effect was observed in graphene quantum dot structure fabricated by e-beam lithography, in which a graphene island is functioned as a quantum dot, while the bulk graphene served as electrodes and the graphene necks with reduced width in between served as the tunneling barriers\textsuperscript{27-31}. Similar to this quantum dot structure formed by width tailoring, GNRs with large edge roughness may also undergo such dot-neck segregation. Therefore, GNRs can be modeled as multiple graphene quantum dots in series. And multiple Coulomb blockade effect was expected at low temperature. Castro Neto and co-workers calculated the transport behavior of edge disordered GNRs due to Coulomb blockade effect\textsuperscript{32}. Their result indicated that the Coulomb charging energy could open up an energy gap at low temperature. The Coulomb gap size scaled with ribbon width with an expression of $E_g = e^2/W \times \exp(-W/W_0)$, $W_0 \approx \pi c/4k_F \tau$ ($\tau$ is the average transmission per channel), which was in good agreement with experimental observation within the error range\textsuperscript{15}.

The low temperature non-linear transport measurements also proved the existence of Coulomb blockade effect in GNR transport properties\textsuperscript{33-36}. Figure 2-3 a and b showed the differential conductance as a function of gate and bias voltage for a GNR device made by nanowire masking method\textsuperscript{37}. The semiconducting diamonds of suppressed conductance was showed to
comprise of numbers of small diamonds, which is quite similar to multiple Coulomb blockade effect$^{38-40}$. Kim and co-worker studied the temperature dependent transport of top gated GNRs$^{36}$. By varying the gate capacity coupling, their result indicated that the Coulomb charging energy contributed at least 60% of the total charge transport activation energy$^{36}$. Additionally, Ihn and coworkers investigated the capacitance coupling of a graphene quantum dot single electron transistor (SET) to an adjacent GNR. Discontinuities in the SET Coulomb resonance were observed as they cross the conductance resonances in the nanoribbons$^{33}$, which clearly indicated charge accumulation event in the GNRs as a function of gate voltage, providing solid evidence for Coulomb charging effect in GNR transport.

On the other hand, Goldhaber-Gordon and co-workers argued that the main origin of the quantum dots formation did not come from edge roughness. Instead, the primary structure of quantum dots were formed by the electron-hole puddles which were induced by charge impurity disorder potentials and separated by confinement gaps (Figure 2-4). They found that the extent of transport gaps did not necessarily depend on ribbon length, and that annealing at elevated temperature tended to shift the charge neutrality point close to zero in companions with reducing transport gaps. The annealing process could reduce the disorder potentials by eliminating the charge impurities, but the temperature was not high enough to affect the edge configuration. Although earlier work demonstrated the length dependent of transport gaps$^{41}$, this should be reexamined because longer ribbon length would definitely lead to more potential fluctuations along the ribbon at fixed impurity concentration. The lack of length dependence found in the annealing study was contradictory to the models of Anderson insulators and edge roughness induced many body effects. In addition, Coulomb blockade was observed in GNR p-n-p junctions created by top-gating or side gating$^{42,43}$. The results pointed out that quantum dots could be induced by electrical
potential variations and further isolated by confinement band gaps. These results strongly supported the model that transport along most experimentally fabricated GNRs appeared to be a tunneling process through multiple quantum dots (electron-hole puddles separated by confinement gaps) due to disorder potentials induced by charge impurities.

To date, the real origin of the transport gap formation is still under debate. In a disorder system such as GNRs, Anderson localization can manifest itself as disorder potential induced back scattering causing vanishing of transport probability. The localization effect reaches its maximum in the vicinity of Dirac point $^{20,24}$, which can certainly contribute to the tunneling barrier in the physical picture of multiple quantum dot models. Also, the temperature dependence of disorder GNRs shows nearest neighbor hopping behavior at elevated temperature and variable range hopping at low temperature $^{36}$, consistent with the description of hopping transport considering both Anderson localization and Coulomb interaction $^{44}$. Therefore, the formation of band gap in real GNR device may involve complex scheme of Anderson localization and Coulomb charging effect. Although GNRs were predicted to have novel electronic and magnetic properties, the real device performances were much more complicated due to edge disorder and charge impurities. Further optimization of fabrication process is necessary in order to control the landscape of charge disorder potential and to reduce the edge roughness.

D. Magnetotransport of GNRs- discovery of very large magnetoresistance

D1. Theoretical study of magnetism in GNRs

As mentioned in introduction, armchair GNRs can develop a direct band gap and zigzag GNRs is predicted to be metallic $^{45-49}$. Indeed, the zigzag GNRs is of more interesting because of the prediction of spin polarized edge state. The localized spin momentums are parallel along the zigzag edge and antiparallel between the two edges with 0 net magnetic momentums $^{49-51}$. This
antiferromagnetic ground state induced energy gradient across the width of the ribbon, caused additional electron-electron interaction. Similar to the electrical structures and band opening in boron-nitride, the potential difference between neighboring carbon atoms in the width direction opens a band gap, rendering semiconducting nature of zigzag graphene at low energy state (Figure 2-5) 49. Similar to armchair GNRs, the band gaps also scales inverse proportional with the ribbon width. Additionally, unlike the antiferromagnetic ground state, the higher energy ferromagnetic state is metallic with parallel aligned magnetic momentums between opposite edges. A possible semiconductor-metal transition is predicted by switching between the antiferromagnetic ground state to ferromagnetic state with external magnetic field, resulted in a very large magnetoresistance52.

The prediction of exceptional magnetism in zigzag GNRs makes it a very interesting system for magnetic sensing and spintronic application. The pioneer idea was proposed by Son et. al. that a half-metallic state could be induced in zigzag GNRs with applied in plain electric field, in which the ribbon behaved as conductor for one spin state and insulator for the other spin state 53. Similar half metal property was introduced by using defects or impurity engineering to selectively suppress electron conduction along one edge54-57. Devices with magnetic domain wall structures by coupling with ferromagnetic electrodes were also proposed to achieve large magnetoresistance. However, owing to the difficulty to achieve atomic control of edge configuration, previous efforts are still limited to theoretical studies.

In the following section, I will present our experimental observation of a significant enhancement in the conductance of a graphene nanoribbon field-effect transistor in a perpendicular magnetic field. A negative magnetoresistance of nearly 100% was observed at low temperatures, with more than 50% remaining at room temperature. This magnetoresistance can be tuned by
varying the gate or source-drain bias. We also find that the charge transport in the nanoribbons is not significantly modified by an in-plane magnetic field. The large values of the magnetoresistance we observe may be attributed to the reduction of quantum confinement by the formation of cyclotron orbits and the possible delocalization effect under the perpendicular magnetic field.

D2. Experimental approach

The graphene nanoribbon-FETs were fabricated using SiO$_2$ nanowires as the physical etching masks. Silicon nanowires with diameters ranging from 5 to 40nm were grown using catalytic chemical vapour deposition, and oxidized in air at 900°C for 10 to 15 min to produce a SiO$_2$ insulating layer or fully oxidized nanowires. Graphene flakes were mechanical peeled from natural graphite onto heavily p-type doped silicon wafers with 300 nm thermal oxide. After that, we physically transferred the SiO$_2$ nanowires from the growth substrate to the graphene substrate via a contact printing approach. Specifically, a graphene device substrate was first firmly attached to a bench top, and the nanowire substrate was placed upside down on top of the graphene device substrate such that the nanowires were in contact with the graphene. A gentle manual pressure was then applied from the top followed by slightly sliding the growth substrate. The nanowires were transferred onto graphene device substrate by the sheer forces during the sliding process. The sample was then rinsed with isopropanol followed by nitrogen blow-dry, in which capillary drying process near the nanowires can help them to firmly attach to the graphene flakes. The nanowire position was determined by dark field microscopy or atomic force microscopy. E-beam lithography was used to define source and drain electrodes according to the position of the nanowires on the graphene, and Ti/Au 7 nm/90 nm film was evaporated with an e-beam evaporator in high vacuum. The graphene regions not protected by the nanowires were etched away by oxygen plasma under 30-40 W for 20s. Transport measurements were carried out in a pumped
D3. Result

The room temperature transport has already been discussed previous in introduction. Unlike conventional e-beam lithography defined graphene nanoribbons in which polymeric resist residue may heavily dope the nanoribbons and results in large positive shift of charge neutrality point, the graphene nanoribbons obtained with our method show relatively neat performance with charge neutrality points typically in the range of 0-5 V in the back-gate configuration.

The electrical transport characteristics of the graphene nanoribbon devices were typically carried out at 1.6 K unless mentioned otherwise. The black line in Figure 2-6a shows a differential conductance with respect to the gate voltage for a typical nanoribbon device with channel width of ~15 nm and length of 800 nm. The curve indicates a strong suppression of conduction in this relatively long device with a transport gap in the gate region of 0.4 to 6.6 V. Figure 2-6b further shows differential conductance as a function of both the gate voltage and the source-drain bias; a diamond like characteristics of suppressed conductance consisting of a number of sub-diamonds is clearly seen. In particular, smaller diamonds are also observed away from the main transport gap region (Figure 2-6b), indicating that the charge transport in our device is related to the model of multiple graphene QDs in series along the nanoribbon. The formation of the QD structure in the nanoribbons may be attributed to edge roughness or local potential variation.

We also performed the magneto-transport measurements under a magnetic field up to 8 T normal to the device plane. Previous magnetoresistance study of large graphene flakes showed a
non-saturated positive magnetoresistance near the minimum conductance point, representing carrier transport through inhomogeneously distributed electron and hole puddles of equal mobility\textsuperscript{60,61}. In contrast, our graphene nanoribbon device exhibits very large negative magnetoresistance that is highly dependent on the exact gate voltage and source-drain bias. Upon applying a magnetic field, the overall conductance increases dramatically with a much reduced transport gap in the gate sweep (red curve in Figure 2-6a). Near the edge of the original transport gap (back gate voltage ($V_g$) range 0.4~1.9 V and 4.6~6.6 V), the differential conductance is essentially switched on from a completely off state upon applying a magnetic field (8 T) with a differential conductance increase up to 1000 fold or more. On the other hand, the average conductance rises about 2-4 times when the device is gated far away from the transport gap region ($V_g < 0.4$ V or $V_g > +6.6$ V). This phenomenon is more evident in the two dimensional differential conductance plots as shown in Figure 2-6 b-d: the overall diamond of suppressed conductance shrinks significantly in both the gate and source-drain bias directions, and those sub-diamonds at the edge of the transport gap region become so conductive that the transport gap reduces to 1.9-4.6 V in gate sweep at 8 T.

Figure 2-7a further shows the current-voltage (I-V) characteristics near the charge neutrality point at different magnetic fields. The I-V curves exhibit non-linear behaviour near zero source-drain bias, in which the semiconducting like nonlinear gaps denoted as source-drain gap decrease from 25 meV at 0 T to 4.3 meV at 8 T (Figure 2-7a inset). Note that the source-drain gap is defined by a steep increase of current in logarithmic scale\textsuperscript{36}. These results suggest that the transport barrier decreases with the increase of the magnetic field. At the edge of the transport gap (Figure 2-6b, $V_g=1$ V) where the sub-diamonds of the suppressed conductance nearly disappear at high magnetic field, the source-drain conduction gap decreases from 16.6 to 0.5 meV (Figure 2-7b
inset). Further away from the charge neutrality point ($V_g=0$ V), the magnetic field has less effect on the device current, although significant change can still be observed, for example, near zero source-drain bias in the small source-drain gap region, which almost totally disappeared at 8 T (Figure 2-7c and inset).

The I-V characteristics indicate that the huge magnetoresistance can be readily obtained by tuning the exact electronic states of the device. Figure 2-7d plots the current ratio of $I(8T)/I(0T)$ versus source-drain bias at $V_g = 0$, 1 and 3 V. All three curves show large increase of current ratio when approaching the transport blockade region. Significantly, over 4 orders of magnitude of increase in current can be observed at the edge of the blockade. This exceptionally large magnetoresistance cannot be well described using conventional formula: magnetoresistance = $\Delta R/R_0$ because current ratio of over 100 already gives magnetoresistance of -99%, while 4 orders of magnitude change gives an magnetoresistance value of -99.99%. To better examine the magnitude of the conductance change, we simply use current (conductance) ratio to describe the exceptionally large magnetoresistance in our device. Figure 2-7d clearly shows that the current ratio can be tuned from a few times up to $>10,000$ times, depending on the exact gate or source-drain bias. Similar scenarios were also observed in the electron-transport branch (Figure 2-8). Additionally, we have fabricated and studied more than ten devices, and all of them exhibit similar negative magnetoresistance.

Figure 2-7e provides a general view of current ratio as a function of both the source-drain bias and gate voltage, which further demonstrates the tunability of the magnetoresistance observed in our graphene nanoribbon devices. In general, the current ratio increases significantly when the device is tuned to the proximity of the diamond blockade region and reaches the highest value at the edge of the conductance suppressing diamond. Within the diamond, the magnetoresistance
value cannot be accurately determined because the current through the device is below our equipment measurement capability. Figure 2-7f shows the magneto-response at gate voltage of 1 V at different source-drain bias. In general, the current ratios increase more rapidly with magnetic field when biased closer to the blockade region. In all cases, the current increases exponentially and shows no evidence of saturation up to 8 T (Note that the figure is in exponential scale).

Electric field control of magnetoresistance has recently attracted considerable attention in multifunctional logic devices. Several material systems, including carbon nanotubes, semiconductor quantum dots and nanowires coupled with ferromagnetic (FM) electrodes, have been explored in this regard but only with limited tunability achieved to date. Importantly, our studies demonstrate that graphene nanoribbons themselves without artificially engineered ferromagnetic contacts can exhibit extraordinary magnetoresistance (nearly -100% with current ratio over 10,000) that is highly tunable by either varying the gate voltage or source-drain bias, thus enabling an entirely new material system and device structure for multifunctional magnetic logic device.

We have also studied the magnetoresistance evolution with increasing temperature. Figure 2-9a and b show the current ratio and negative magnetoresistance versus source-drain bias at variable temperatures at $V_g = 3$ V. The maximum current ratio decreases from more than 10,000 at 1.6 K, to nearly 20 (magnetoresistance = -95%) at 25 K, and to 3 (magnetoresistance = -70%) at 77 K, as the conductance suppression due to conduction band gap and/or Coulomb blockade effect is weakening with the increase of temperature. The conduction band gap apparently disappears at room temperature as the device shows linear transport behaviour (Figure 2-9c). At this point, the magnetoresistance can no longer be modified by source-drain bias voltage, and is also only weakly tunable by gate voltage. Nonetheless, the negative sign of magnetoresistance persists up to room
temperature. A nearly linear increase of magnetoresistance with magnetic field and up to -56 % magnetoresistance is obtained at 8 T at $V_g = 3 \text{ V}$ (inset of Figure 2-9c). For practical considerations, it is valuable to note the magnetoresistance at low magnetic field. The room temperature magnetoresistance reaches ~ 4% at a low magnetic field of 0.5 T, which is not as striking as low temperature data (up to 50 % at 0.5 T), but is still very significant in a device with no ferromagnetic materials. It is now well known that the conduction band gap of graphene nanoribbon is inversely proportional to the ribbon width, and sub-5 nm nanoribbons can develop large enough conduction band gap to completely shut off the conductance even at room temperature. Based on our observation of magnetoresistance enhancement near the conductance suppression diamond, larger magnetoresistance at room temperature may be achievable when ribbon width is further shrunk down to the sub-5 nm regime.

To further elucidate the magnetoresistance effect in our device, we have fabricated shorter channel graphene nanoribbon device with channel length of 200 nm, aiming to reduce the number of QDs along the nanoribbon device$^{35}$. Indeed, simpler transport characteristics with better resolved Coulomb blockade structures were observed in this shorter device (Figure 2-10a), although the jointed diamonds still indicated multiple coulomb blockade effect$^{64}$. Upon applying a magnetic field of 3 T and 6 T, the shrinkage of each diamond is clearly seen (Figure 2-10b, c). The size of the diamond structure (bias gap) shrinks consistently with increasing magnetic field (Figure 2-11). The gate sweeps show overall conductance increase: the conductance peak grows and broadens; and conductance valley arises, resulting in a diminishing of the blockade region (Figure 2-10d). Interestingly, the blockade position is not significantly shifted with magnetic field up to 3 T, indicating no significant change of QD configuration. These diamonds almost disappear when
the magnetic field is increased to 6 T, suggesting that charge hopping through the nanoribbon can be significantly enhanced under an external magnetic field.

D4. Discussion

Previous theoretical studies have predicted interesting magneto-transport properties in graphene nanoribbons with multiple possible origins. For example, first-principle calculations have predicted the existence of semiconducting antiferromagnetic spin state in the zigzag graphene nanoribbons that can be excited to metallic ferromagnetic state with an applied magnetic field. However, this possibility was eliminated in our devices by performing magnetoresistance measurement with an in-plane magnetic field in which no obvious magnetic response was found (Figure 2-12). On the other hand, recent theoretical studies indicate that a perpendicular magnetic field can greatly modify lateral quantum confinements in graphene nanoribbons and profoundly impact the charge transport due to the formation of cyclotron orbits originated from Dirac-Landau-level behaviour in the graphene nanoribbons. It is suggested that a perpendicular magnetic field can induce cyclotron orbits of the electrons motion with the cyclotron length $l_B = (\hbar/eB)^{1/2}$, where $e$ is the electron charges and $B$ is magnetic field. At lower magnetic field with large cyclotron length, the cyclotron wavefunction extends beyond the ribbon width, electron motion is strongly affected by the ribbon edges and the quantum confinement dictates the electrical properties of the graphene nanoribbons. With increasing magnetic field, the cyclotron length decreases. When the cyclotron length becomes comparable or smaller than ribbon width, the ribbon edges and confinement potential become less relevant in the ribbon electrical properties, and it would eventually completely eliminate the edge confinement to induce a semiconductor-metal transition at high enough magnetic field. Recent tight-binding calculations have shown the confinement band gap of semiconducting armchair graphene nanoribbons indeed shrinks continuously with
increasing magnetic field\textsuperscript{67}. The cyclotron length $l_B = (\hbar/eB)^{1/2}$ is 26.6 nm at 1 T and 9.4 nm at 8 T, comparable to our nanoribbon widths, suggesting that the magnetic field can indeed significantly modify the electronic structure and charge transport in our graphene nanoribbon devices.

In a typical graphene nanoribbon with multi-coulomb blockade transport characteristics due to experimental edge roughness and local potential variation, the confinement gaps function as energy barriers for charge transport across different electron hole puddles along the nanoribbon\textsuperscript{28,35}. Therefore, a decrease of confinement gap with magnetic field can reduce the charge hopping barrier in the multi-coulomb blockade device, and effectively reduce the overall conduction band gap with increasing magnetic field, as we have observed in our devices. This is also consistent with the evolution of diamond structure in our short channel device, in which the sizes of the Coulomb diamonds clearly shrink with increasing magnetic field.

Our temperature dependent studies also show that the thermal activation energy of charge transport in the graphene nanoribbons decreases with increasing magnetic field, further suggesting the shrinkage of the confinement gaps by the perpendicular magnetic field\textsuperscript{44}. As shown in Figure 2-13, the dash line at high temperature region of each magnetic field is a fit to simple thermal activated transport: $G_{\text{min}} \sim \exp(-E_a/2k_BT)$; and the dash line at low temperature region is a fit to variable range hopping: $G_{\text{min}} \sim \exp(-(T_0/T)\gamma)$ with $\gamma = 0.4$. Here we studied the temperature dependent charge transport in the transport gap region. The minimum conductance (off-resonance differential conductance near the charge neutrality point) was plotted with temperature at different magnetic field\textsuperscript{36}. The plot shows that high temperature charge transport follow thermal activated behavior given by $G_{\text{min}} \sim \exp(-E_a/2k_BT)$. At low temperature region, the transport can be described by quasi-1D variable hopping behaviour with $G_{\text{min}} \sim \exp(-(T_0/T)\gamma)$, where $\gamma$ is a dimensional factor in the range of $1/2\sim 1/3\textsuperscript{70}$. The high temperature activation energy $E_a$ can be obtained by linear
fitting of the Arrhenius plot (dash line), and $E_a$ is 291 KT (25.1 meV) at 0T for this particular device. Interestingly, the activation energy decrease with increasing vertical magnetic field: the value drops to 196 KT (16.9 meV) at 4T and further reduces to 145 KT (12.5 meV) at 8T. Previous theoretical studies show that the high temperature $E_a$ contributed by edge roughness and Coulomb interaction can be approximated by $(0.2t+t/\varepsilon)/W$, where confinement gap $E_g \sim t/W$ ($t$ is the nearest hopping element, $\varepsilon$ is the dielectric constant of the embedded medium and $W$ represent the width)\cite{44}. Therefore, the decrease of high temperature activation energy with magnetic field indicates the shrinkage of the confinement gaps by modifying the hopping matrix with additional magnetic flux\cite{67}.

Additionally, the edge roughness induced back scattering in terms of strong localization may also contribute to the observed conductance suppression\cite{19}, in which magnetic field can cause the delocalization effect and suppress the back scattering\cite{71}. The Coulomb blockade effect can significantly enhance the magnetoresistance at the edge of the transport gap/source-drain gap where the charge transport are more sensitive to the magnetic field. On the other hand, when gated or biased outside the transport gap or under high temperature in which current flow is dominated by drifting and the conductance is limited by edge or charge impurity scattering, the observed magnetoresistance is expectedly not as striking as that under tunnelling condition.

**E. Summary**

An extraordinarily large tunable magnetoresistance is demonstrated in graphene nanoribbon-FET devices. Negative magnetoresistance nearly 100% with over 10,000 times conductance increase was demonstrated at 1.6 K and negative magnetoresistance nearly 56 % was obtained at room temperature. This magnetoresistance can be readily tuned by gate voltage and source-drain bias in which the enhancement reaches the maxima near the edge of the conduction
band gap. Although further experimental and theoretical studies will be necessary to fully elucidate the exact mechanism responsible for the observed negative magnetoresistance, our experimental findings clearly demonstrate that the graphene nanoribbons exhibit interesting magneto-transport properties and may open exciting opportunities in magnetic sensing and a new generation of magneto-electronic devices.
Reference:

Figure 2-1 | Experimental measured conduction band gaps $E_g$ as a function of ribbon widths with different sets of crystal orientation. The inset shows $E_g$ vs relative angles for the device sets of different orientation. Dashed lines in the inset show the value of $E_g$ as predicted by the empirical scaling of $E_g$ vs W. (Adapted from [15])
**Figure 2-2** | Anderson localization induced conductance suppression in edge disordered AGNRs.  
**a,** Illustration of armchair nanoribbons with edge disorder. Simulated conductance with edge disorder probability $P=0.3$ **b** and 0.05 **c.** Doted and dashed lines indicate edge disordered sample with different length (in unit cell), compare to nanoribbons without edge defects (solid).  
(Adapted from [19])
Figure 2-3| Schematic of quantum dots formation in GNRs due to disordered potentials and confinement gaps. (Adapted from [35])
Figure 2-4: Experimental demonstration of Coulomb blockade diamond like transport features in GNR FETs. 

**a, b,** Differential conductance as a function of source-drain bias and gate voltage with channel length 1μm and width ~ 30 nm.

**c,** Zero source-drain bias differential conductance vs gate voltage. All measurements were done at 1.6K. (Adapted from [37])
Figure 2-5: Contour graph for localized edge spin state in zigzag GNR. 

**a,** Contour graph for localized edge spin state in zigzag GNR. The blue and red colors indicate different spin directions.

**b,** Corresponding band structure. The arrows indicate direct band gap $\Delta_z^0$ and energy splitting $\Delta_z^1$ at $kd_z=\pi$.

**c,** The evolution of energy gaps with ribbon width. (Adapted from [49])
Figure 2-6| Electrical transport measurement of a graphene nanoribbon-FET with width of ~15 nm and length of 800 nm. a, Differential conductance versus gate voltage with a magnetic field of 0 T (black) and 8 T (red) normal to the device plane. The measurements were carried out at 1.6 K. b-d, Differential conductance as a function of source-drain bias and back-gate voltage under magnetic field of 0 T (b), 2 T (c), and 8 T (d). These measurements show diamonds of suppressed conductance shrunk both in source-drain bias and gate voltage direction with increasing magnetic field.
Figure 2-7| Tunable magnetoresistance in graphene nanoribbon-FET. a-c, The impact of magnetic field on current-voltage characteristics when the device is gated at V=3 V (a), 1 V (b) and 0 V (c). Each inset shows the source-drain gap (ΔVsd) as a function of magnetic field. d, Current ratio I(8T)/I(0T) versus source-drain bias at Vg = 3 V. The middle interval for each plot is in the range of suppressed conductance which is beyond our equipment measurement limits. e, Current ratio I(8T)/I(0T) as a function of source-drain bias and gate voltage, highlighting huge increase of current under magnetic field when probing the device close to the diamond of suppressed conductance. f, Current ratio I(M)/I(0T) as a function of magnetic field when source-drain is biased at -10 mV, -20 mV, -25 mV and -45 mV at Vg = 1 V.
Figure 2-8 | Current ratio I(8T)/I(0T) as a function of source-drain voltage when gated at \( V_g = 3.0 \) V, 4.7 V and 6 V, highlighting that the significant magnetoresistance (MR) can also be obtained at electron-transport branch.
Figure 2-9 | Temperature dependent magneto-transport properties. a-b, Current ratio $I(8T)/I(0T)$ (a) and negative magnetoresistance (MR) (b) as a function of source-drain bias at 1.6 K, 5 K, 25 K, 77 K and 285 K. The device was gated at 3 V. c, Room temperature (285 K) I-V characteristic ($V_g=3$ V) at different magnetic field. The inset shows the negative magnetoresistance (MR) increases linearly with the applied magnetic field.
Figure 2-10| Magneto-transport properties of a short channel graphene nanoribbon-FET device with width 37 nm and length 200 nm. a-c, Differential conductance measurements as a function of source-drain bias and gate voltage at (a) 0 T, (b) T and (c) 6 T showing the evolution of diamond of suppressed conductance region with increasing magnetic field. d, Differential conductance (G) versus gate voltage at 0 source-drain bias at magnetic field of 0 T, 3 T and 6 T.
Figure 2-11| Evolution of source-drain gap at selected gate voltages for the short channel graphene nanoribbon FET device shown in Figure 2-10.

Figure 2-12| Current-voltage characteristics under in-plane magnetic field. No obvious conductance change is observed with in-plane magnetic field up to 8 T, in contrast to the case with perpendicular magnetic field described in the main article. This observation excludes the magnetoresistance origination from magnetic edge states which have small magnetic field angular dependence due to weak spin-orbital coupling.
Figure 2-13 | Temperature dependence of the minimum conductance at different magnetic field.
A. Introduction to graphene RF transistors

Graphene is of considerable interest as a potential new electronic material. In particular, it has attracted enormous attention for radio frequency transistor applications owing to its exceptional high carrier mobility, high carrier saturation velocity and large current density. The observed charge carrier mobility can reach 200,000 cm$^2$V$^{-1}$s$^{-1}$ for both electrons and holes even at room temperature. Unlike some other high mobility semiconductors, such as InSb which can only exhibit high mobility (mobility exceeding 60,000 cm$^2$V$^{-1}$s$^{-1}$ at room temperature) at a very low doping level with a 2D electron gas carrier concentration of $10^{11}$ cm$^{-2}$, graphene retains high mobility even with electron or hole doping with carrier density > $10^{12}$ cm$^{-2}$, enabling high current delivery through the device. The observation of quantum hall effect at room temperature also indicated exceptional electronic quality of graphene crystals. These extraordinary electrical properties make graphene a good candidate for high frequency device applications.

B. Dielectric integration of graphene

However, the fabrication of high performance graphene transistors is of significant challenge since the conventional device fabrication process steps can often introduce undesired damage into graphene lattice to degrade its electronic performance or result in non-ideal device geometry with excessive parasitic capacitance or serial resistance. The recent development of self-aligned graphene transistors with a nanowire gate can address these challenges, and has enabled graphene transistors with the highest intrinsic cut-off frequency up to 300 GHz. Moving forward to the terahertz regime requires high quality graphene material, damage-free dielectric integration strategy, and self-aligned device layout. The strategy of physical assembly of nanowire gate is promising for addressing the last two problems. On the other hand, the scalability of this approach
is complicated with the requirement of unconventional nanowire assembly processes. Instead of fighting with the difficulty of nanowire assembly, here we report a scalable approach to high performance graphene transistors by transferring lithographically patterned gate stacks onto graphene as the self-aligned top-gate to demonstrate the highest cut-off frequency up to 427 GHz.

C. Fabrication of self-aligned graphene transistors with transferred gate stacks

Figure 3-1 illustrates our approach to fabricate the self-aligned graphene transistors with transferred gate stacks. A 50-nm gold thin film is first deposited on a Si/SiO$_2$ substrate by e-beam evaporation. This gold film is served as the sacrificial substrate in the transferring process. Subsequently, we build the (Al$_2$O$_3$/Ti/Au) gate stacks on top of the gold film by standard atomic layer deposition (ALD), lithography and reactive ion etch (RIE) processes (Fig. 3-1a, b). The gate sidewall spacer is formed by depositing a thin layer Al$_2$O$_3$ film using ALD approach (Fig. 3-1c), followed by an anisotropic RIE process to etch away unwanted Al$_2$O$_3$ film on the top surface of the gate metal and the substrate (Fig. 3-1d). Because of the limited affinity between the sacrificial Au film and the underneath SiO$_2$, the fabricated gate stacks are easily peeled off by thermal release tape. Additionally, to facilitate peeling off and releasing, a thin layer of polymer which has glass transition temperature close to the thermal tape releasing temperature is spin-casted before peeling off the gate stacks. Therefore, the patterned top-gate stacks are sandwiched between the gold layer and the polymer layer for peeling off from the Si wafer (Fig. 3-1e). After Au etching, the gate stacks can be readily transferred onto desired graphene substrate through a thermal releasing process (Fig. 3-1f). Compared to other releasing method such as lateral undercut etching of the underneath sacrificial layer, mechanical peeling is much faster and has lower chemical degradation or contamination to the releasing structure. Moreover, the flexible nature of our approach is compatible with roll-to-roll transfer process, making it highly potential for printable electronics.$^{21}$
The CVD graphene is grown on copper foil and transferred onto arbitrary substrates such as silicon, glass and plastic\textsuperscript{22,23}. After transferring the gate stacks, external source, drain and gate connections are formed by conventional lithography and metallization process. Finally, a thin layer of Pd/Au (5 nm/10 nm) is used to form the self-aligned source and drain electrodes, which is separated by the gate spacer dielectrics (Fig. 3-1h, i).

This fabrication approach is intrinsically scalable, and compatible to various substrates such as Si, glass and plastics. As shown in Figure 3-2, arrays of self-aligned graphene transistors with uniform device geometry and spacing are demonstrated on glass (Fig. 3-2a) and 300 nm Si/SiO\textsubscript{2} substrate (Fig. 3-2b). Figure 3-2c shows a scanning electron microscopy (SEM) image of the top view of an individual self-aligned graphene transistor. The cross-sectional transmission electron microscope (TEM) image shows that the self-aligned source and drain electrodes are well separated and precisely positioned next to the gate spacer dielectrics (Fig. 3-2d). Since the device performance is essentially determined by the structure of the individual gate stack and the self-aligned source drain electrodes, the relative large gap between external source-drain lead electrodes may not affect the device performance, rather making it more tolerant to the transfer induced misalignment and distortion of the gate stacks arrangement.

The conventional dielectric integration approach can often introduce substantial defects into graphene lattice and degrade its electronic performance (e.g. carrier mobility). To investigate the impact of our transferring method on graphene carrier mobility, we have studied the electrical performance and carrier mobility distribution of more than 20 graphene transistors (in the back-gate configuration) before and after dielectric integration. Here the CVD grown graphene is patterned by photolithography with a channel length \( \approx 8 \mu \text{m} \) and a channel width \( \approx 8 \mu \text{m} \). The effective mobility values are extracted from the drain-source current (\( I_{ds} \)) versus the back-gated...
voltage \((V_{BG})\) (\(I_{ds}-V_{BG}\)) curve. Importantly, a histogram of the mobility value shows that the CVD grown graphene exhibits a similar mobility distribution in the range of 1000-2000 cm\(^2\)/V·s before and after the transfer dielectric integration process (Fig. 3-3a). These studies clearly demonstrate our transfer gate approach does not lead to an obvious degradation of the electronic performance of the graphene.

Figure 3-3b-f depicts the room-temperature electrical transport characteristics of the self-aligned graphene transistors. Before characterizing the transistor properties of our self-aligned devices, gate leakage current (\(I_{gs}\)) from the gate stack to the underlying graphene is measured from -4 V to 4 V gate voltage, which indicates the gate leakage current is negligible during the measurement. Figure 3-3b shows the \(I_{ds}-V_{ds}\) output characteristic of a 300 nm channel length self-aligned graphene transistor at various gate voltages. The maximum scaled on-current of 1.73 mA\(\mu\)m\(^{-1}\) can be achieved at \(V_{ds} = -1\) V with slight current saturation. The current saturation is desirable for the power gain performance in radio frequency (rf) graphene transistors (10).

The \(I_{ds}-V_{TG}\) curve of the same device is measured at different drain bias from 100 mV to 600 mV with the top-gate voltage sweeps from 0 V to 3 V (Fig. 3-3c). It shows a typical characteristic of p-type doping with the Dirac points located at positive top-gate voltage and can be attributed to oxygen doping occurred during the growth or transfer processes\(^{10}\). Overall, the hole-transport branch can switch from saturation current to Dirac point within 2 V of top-gate voltage sweeping, indicating a strong top gate capacitance coupling. In general, these top-gated self-aligned graphene transistors exhibit a very small \(I_{ds}-V_{TG}\) hysteresis \(\sim 0.1\) V or less, demonstrating the relatively clean nature of the graphene–dielectric interface. A suppression of electron transport branch is observed in the \(I_{ds}-V_{TG}\) transfer curves. This electron-hole asymmetry is commonly originated from the misalignment between the work function of contact electrodes and the neutrality point of the
A trend of negative shift of Dirac point is observed with increasing drain voltage. This can be explained by the fact that the Dirac point will shift by $1/2 \Delta V_{ds}$ with a change of bias voltage $\Delta V_{ds}$ due to the relative potential between the gate and drain\cite{25,26}.

In order to characterize the gate capacitance, the conductance of one graphene transistor with 300 nm channel length and 22 nm dielectric thickness is measured as a function of both top-gate voltage ($V_{TG}$) and back gate voltage ($V_{BG}$) (Fig. 3-3d). The ratio between the top-gate and back-gate capacitance is extracted from the slope of the linear shift trajectory of the Dirac point as a function of both the top-gate and back-gate voltage, which gives a value of $C_{TG}/C_{BG}$ about 29. For 300 nm SiO$_2$, the back-gate capacitance is $C_{BG} = 11.5$ nF cm$^{-2}$, therefore, the estimated top-gate capacitance is $C_{TG} = 334$ nF cm$^{-2}$ (ref. 27), which is consistent with the result obtained from geometry based finite-element calculations ($C_{TG} = 359$ nF cm$^{-2}$).

Figure 3-3e shows the $I_{ds}$-$V_{TG}$ transfer curves of several devices with variable self-aligned gate length ranging from 3 $\mu$m to 100 nm. With decreasing channel length, a general trend of positive shift of Dirac point and decrease of on/off ratio is observed, which can be explained by short channel effect\cite{25}: in short channel device, the off-state energy barrier is strongly affected by drain voltage, it thus increases the off-state current and requires higher gate voltage to turn off the channel. Figure 3-3f, shows the extracted transconductance, $g_m = \left| \frac{dI_{ds}}{dV_{TG}} \right|$, for devices with different channel length. The peak transconductance at bias of 600 mV increases from 0.11 mS/$\mu$m ($L=3\mu$m) to 0.53 mS/$\mu$m ($L=300$ nm). But a further shrinkage of channel length to 100 nm leads to a reduction of transconductance to 0.45 mS/$\mu$m. It may be explained by the short channel effect with less effective gate modulation\cite{25}. Moreover, the occurrence of Klein tunneling in short channel device makes the short-channel effect worse\cite{28-31}.

D. RF performance of self-aligned graphene transistors with transferred gate stacks
The above discussion clearly demonstrates that our self-aligned graphene transistors with transferred gate stacks exhibit excellent d.c. performance. To further determine the cut-off frequency \( f_T \) of our devices, we carried out the on-chip microwave measurements with an Agilent 8361A network analyzer in the range of 50 MHz - 30 GHz. The graphene transistors for rf measurement are fabricated on glass substrate in order to minimize the parasitic pad capacitance. To accurately assess the intrinsic \( f_T \) value, careful de-embedding procedures are performed using the exact pad layout as “open”, “short” and "through" structures on the same chip. The de-embedded S parameters constitute a complete set of coefficients to describe intrinsic input and output behavior of graphene transistors.

Figure 3-4a shows the small signal current gain \(|h_{21}|\) extracted from the measured S parameters at \( V_{TG} = 1.5 \) V and \( V_{ds} = 0.6 \) V in a 220 nm channel length graphene transistor. The curve shows a typical \( 1/f \) frequency dependence expected for an ideal FET. The linear fit yields an \( f_T \) value of 57 GHz for this particular device (Fig. 3-4a), which is also verified by using Gummel’s approach (Fig. 3-4a, inset). To further probe the limit of the frequency response, we have fabricated graphene transistors with smaller channel lengths. Figure 3-4b, c show the result extracted from another two self-aligned graphene transistors with 100 nm and 46 nm channel length. The cut-off frequency is \( f_T = 110 \) GHz and \( f_T = 212 \) GHz respectively. After the rf measurement, we have carefully analyzed the S parameters for all three devices. The device component values (including gate-source capacitance, gate-drain capacitance, transconductance, source resistance and drain resistance) derived from the rf measurements are consistent with those obtained from the d.c. measurements and finite element simulations, demonstrating the validity of the rf measurements and the de-embedding procedures.

In addition to \( f_T \), maximum oscillation frequency (\( f_{\text{MAX}} \)), defined as the frequency at which
the power gain is equal to one, is another important parameter for rf characteristics. The power gain performance plot shows a device with 220 nm channel length exhibits a high $f_{\text{MAX}}$ of 29 GHz and a 46 nm device shows a value of 8 GHz. Since $f_{\text{MAX}}$ highly depends on $f_t$, gate resistance and source-drain conductance and does not always scale with channel length. The value of $f_{\text{MAX}}$ of the devices can be further improved by increasing graphene quality, reducing gate resistance and increasing source drain current saturation.

To further investigate the reproducibility of our approach and examine the length-scaling relationship, we examined more than 40 graphene transistors of variable channel lengths (L) and dielectrics thicknesses. In general, the peak cut-off frequencies follow 1/L dependence, which is consistent with previous studies\textsuperscript{16}. Although 1/L\textsuperscript{2} dependence was observed in longer channel devices\textsuperscript{32}, the dominance by contact resistance and degradation of transconductance in short channel transistors leads to 1/L dependence of the cut-off frequency\textsuperscript{16}. For devices with dielectric thickness of 44 nm, the cut-off frequency falls beneath the 1/L trend when the channel length shrinks to 100 nm, which can be improved by using thinner dielectric thickness (e.g., 22 nm and 13 nm) (Fig. 3-4d). This phenomenon can be attributed to the short channel effect\textsuperscript{25}, in which the gate modulation is less effective when the channel length is reduced.

The cut-off frequency of our self-aligned devices shows a significant improvement over previously reported CVD graphene transistors of comparable channel length (e.g. $f_{\text{T}}$~212 GHz for 46 nm device in our approach vs. 155 GHz for 40 nm device reported previously); Nonetheless, the performance of these devices is still far from those obtained from peeled graphene\textsuperscript{14,16}, which suggests that the ultimate performance of our devices here is limited by the quality of CVD graphene rather than the fabrication process. To demonstrate that our approach is applicable for higher performance devices, we have studied the self-aligned transistors on peeled graphene. The
devices are fabricated on highly resistive Si substrate with 300 nm SiO2 due to the difficulty in visualizing the peeled graphene on glass substrate. To accurately assess the intrinsic $f_T$ value, careful de-embedding procedures are performed using the identical pad layout as “open”, “short” and "through" structures on the same chip, following previously established approach\textsuperscript{14}. Importantly, electrical characterization shows graphene transistors with substantially higher cut-off frequency can be obtained in this way. The $I_{ds}$–$V_{TG}$ transfer of a 67 nm device shows that a maximum scaled on-current of 3.56 mA/μm and a peak scaled transconductance of 1.33 mS/μm is obtained at $V_{ds} = 1$ V (Fig. 3-5a). Figure 3-5b shows the small signal current gain $|h_{21}|$ of a 67 nm channel length graphene transistor with a typical $1/f$ frequency dependence and an extracted $f_T$ value of 427 GHz at $V_{ds} = 1.1$ V (Fig. 5B and inset). Additionally, an $f_T$ value of 169 GHz can be obtained at $V_{ds} = 0.4$ V, indicating a linear trend of $f_T$ value with source-drain voltage. The 427 GHz $f_T$ value represents the highest $f_T$ value reported for any graphene transistors to date. We believe the rf performance of our device can be further improved by minimizing the contact resistance or decreasing the gate dielectric thickness to improve gate coupling.

E. Summary

In summary, we have developed a scalable method to fabricate self-aligned graphene transistors on glass with transferred gate stacks. With a damage-free transfer process and self-aligned device structure, the fabricated graphene transistors exhibit the highest cut-off frequency to date in both CVD grown graphene transistors (212 GHz) and peeled graphene transistors (427 GHz). By processing conventional lithography, deposition, etching steps on a sacrificial substrate before integrating with large area CVD grown graphene, this approach defines a pathway to scalable fabrication of high speed self-aligned graphene transistor arrays on arbitrary substrate.
References

9. DOI: 10.1038/nmat3169.
17. DOI: 10.1021/nl201922c.


Figure 3-1 | Schematic illustration of the fabrication of self-aligned graphene transistors with transferred gate stacks. **a**, A 50-nm gold film is first deposited on a Si/SiO$_2$ substrate by $e$-beam evaporation followed by a standard atomic layer deposition (ALD) of Al$_2$O$_3$ film. **b**, Reactive ion etch (RIE) process is employed to pattern the dielectric strips after standard lithography and metallization process. **c**, The gate sidewall spacer is formed by depositing a thin layer Al$_2$O$_3$ film using ALD approach. **d**, An anisotropic RIE process is used to etch away unwanted Al$_2$O$_3$ film on the top surface of the gate metal and the substrate. **e**, A layer of polymer which has glass transition temperature close to the thermal tape releasing temperature is spin-casted before applying thermal releasing tape and peeling off the gate stacking. **f**, The patterned top-gate stacks are peeled off from the Si wafer. After etching away the gold film, the gate stacks can be readily transferred onto desired graphene substrate through a thermal releasing process. **g**, Polymer is removed by an acetone rinse, only leaving the gate stacks on top of graphene strips. **h**, The external source, drain,
and top-gate electrodes are fabricated using electron-beam lithography. Deposition of 5 nm/10 nm Pd/Au metal film to form the self-aligned source and drain electrodes. i, The cross-sectional view of the self-aligned device.

**Figure 3-2** The self-aligned graphene transistor. a, Photo image of large scale self-aligned devices with transferred gate stacks on glass substrate. b, Optical image of self-aligned graphene transistors on 300 nm SiO$_2$/Si substrate. The scale bar is 100 μm. c, The SEM image of a graphene transistors with transferred gate stack. The scale bar is 2 μm. d, Cross-sectional TEM image of the overall device layout. The scale bar is 30 nm.
Figure 3-3| Room-temperature electrical characteristics of the CVD graphene transistors with transferred gate stacks.  

a, The distribution of device mobility before and after the dielectric transfer process.  
b, $I_{ds}$-$V_{ds}$ output characteristics at various gate voltages ($V_{TG}= 0, 1, 1.5, 2.0$ and $2.5$ V) for a $300$ nm channel length self-aligned device.  
c, The transfer characteristics at different bias voltage for the $300$ nm channel length self-aligned device ($V_{ds}= -0.1, -0.2, -0.4$ and $-0.6$ V).  
d, Two-dimensional plot of the device conductance for varying $V_{BG}$ and $V_{TG}$ biases for the self-aligned graphene device.  
e, Transfer characteristics of self-aligned graphene transistor at $V_{ds}= -0.6$ V with channel lengths $3 \mu m$, $1 \mu m$, $300$ nm, and $100$ nm. The channel width is $5 \mu m$ for all devices.  
f, The corresponding transconductance of the devices shown in Fig. 3e at $V_{ds}= -0.6$ V.
Figure 3-4| Radio frequency performance of self-aligned CVD graphene transistors. a-c, Small-signal current gain $|h_{21}|$ versus frequency for three devices with a channel length of (a) 220 nm, (b) 100nm and (c) 46nm at room temperature. The cut-off frequencies are 57 GHz, 110 GHz and 212 GHz respectively at a d.c. bias of 0.6 V. Insets, linear fitting using Gummel’s method, showing extraction of cut-off frequencies identical to the value obtained in the main panel for each device. d, Peak $f_T$ as a function of gate length from over 40 devices with 3 different dielectric thickness.
Figure 3-5| Room-temperature dc and rf characteristics of the self-aligned peeled graphene transistor with transferred gate stacks. a, The transfer characteristics and corresponding transconductance at a d.c. bias voltage of 1 V for the 100 nm channel length self-aligned peeled graphene device. b, Small-signal current gain $|h_{21}|$ versus frequency for the 100 nm peeled graphene device under two different d.c. bias voltage: 1V (solid block) and 0.4 V (hollow block). The inset shows the extraction of $f_T$ by Gummel’s method.

Figure 3-6| The Raman spectrum of CVD grown graphene on SiO$_2$/Si substrate. The ratio of G peak to 2D peak reveals the single layer property of CVD grown graphene.

Figure 3-7| Gate-leakage current versus top-gate voltage ($I_{gs}$ - $V_{TG}$). With the self-aligned
Pd/Au source drain electrodes, the gate-source leakage remains very small compared to the channel current in the range of $V_{TG} = -4$ to $4 \text{ V}$. The leakage current does not significantly affect the transistor characteristics.

Figure 3-8| The back and forth sweep of $I_{ds}$-$V_{TG}$ curve of a peeled graphene device with transferred gate stack. The curve shows a small hysteresis < 0.07 V under ambient conditions at $V_{ds} = 1 \text{ V}$, highlighting the excellent dielectric quality of the gate stack.

Figure 3-9| Finite element simulation of the electrostatic capacitance between transferred gate stack and graphene. The simulated electrostatic capacitance normalized by graphene channel area is about 359 nF/cm$^2$. 
Figure 3-10 | Plot of power gain versus frequency. Mason’s unilateral gain versus frequency for four devices with different channel lengths is plotted. A peak $f_{\text{max}}$ of 29 GHz is obtained from the 220 nm device.

<table>
<thead>
<tr>
<th>Device</th>
<th>Width (μm)</th>
<th>Length (nm)</th>
<th>$g_m$ (mS)</th>
<th>$C_{gs}$ (fF)</th>
<th>$C_{gd}$ (fF)</th>
<th>$R_s$ (Ω)</th>
<th>$R_d$ (Ω)</th>
<th>Projected $f_T$ (GHz)</th>
<th>$f_T$ (GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>220</td>
<td>3.9</td>
<td>8.10</td>
<td>2.86</td>
<td>14</td>
<td>18</td>
<td>56</td>
<td>57</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>100</td>
<td>3.5</td>
<td>3.86</td>
<td>1.03</td>
<td>11</td>
<td>15</td>
<td>114</td>
<td>110</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>46</td>
<td>3.4</td>
<td>2.04</td>
<td>0.41</td>
<td>10</td>
<td>19</td>
<td>221</td>
<td>212</td>
</tr>
</tbody>
</table>

Table 3-1. The component parameter values for three CVD grown graphene devices. $f_T$ is the cut-off frequency with de-embedding. The projected $f_T$ values for all devices with de-embedding process can be determined based on DC device parameters. Also, the critical device parameters (including $C_{gs}$, $C_{gd}$, $g_m$, $R_s$, $R_d$ and $f_T$) derived from S-parameters are consistent with the values determined from DC measurements or electrostatic simulations, demonstrating the validity of the rf measurements and the de-embedding procedures.
Chapter III: BENCHMARKING FEW-LAYER MoS₂ TRANSISTORS AND CIRCUITS FOR HIGH-SPEED FLEXIBLE ELECTRONICS

A. Introduction to transition metal dichalcogenides

Two-dimensional layered materials (2DLMs), such as molybdenum disulfide (MoS₂), are emerging as an exciting material system for future electronics due to their unique electronic properties and atomically thin geometry. Here we report the best performed MoS₂ transistors with optimized device geometry and unprecedented performance, including an on-off ratio exceeding $10^7$, excellent current saturation and a highest intrinsic gain over 30. On-chip microwave measurements demonstrate a highest intrinsic cut-off frequency $f_t$ of 42 GHz and a maximum oscillation frequency $f_{\text{MAX}}$ of 50 GHz, exceeding that of graphene transistors ($f_{\text{MAX}} \sim 44$ GHz). Furthermore, we have, for the first time, demonstrated the integration of multiple MoS₂ transistors on quartz and flexible substrates to form a logic inverter or radio frequency amplifier with voltage gain in the gigahertz regime. This study benchmarks atomically thin MoS₂ for high speed transistors and circuits, and defines exciting potential of 2DLMs for future high speed electronics and flexible electronics.

Two-dimensional layered materials (2DLMs), such as graphene or molybdenum disulfide (MoS₂), are emerging as an exciting material system for future electronics due to their unique electronic properties and atomically thin geometry¹⁻⁸. Graphene has been explored for ultrahigh speed transistors with the intrinsic cut-off frequency exceeding 400 GHz⁹, but typically with insufficient current on-off ratio and little voltage gain due to its zero band gap semimetal nature. Alternatively, the semiconducting MoS₂ have recently drawn considerable interest for overcome these disadvantages of graphene to enable atomically thin transistors with high on-off ratio and intrinsic voltage gain¹⁰⁻¹⁷. A single layer of MoS₂ consists of a layer of Mo atoms sandwiched
between two layers of S atoms. As a 2D material, it shares many interesting characteristics of the well-known graphene such as atomically thin thickness, excellent electronic properties, high mechanical flexibility, and partial optical transparency. With a direct band gap of 1.8 eV, it overcomes the key shortcomings of graphene for electronic applications — the lack of band gap and obvious current saturation. 2D electronics based on single or few-layer MoS₂ represents the ultimate limit of thickness for pushing the limits of the Moore’s Law. With a larger band gap than silicon, MoS₂ is also advantageous for suppressing the source-to-drain tunnelling current in transistors at the scaling limit. Moreover, MoS₂ and other transition metal dichalcogenide (TMD) materials are attractive as an alternative material for low-cost flexible electronics that is currently dominated by amorphous silicon and organic semiconductors with low mobilities around 1 cm²/V·s or less.

A key step to realize the electronic application using 2DLMs is the demonstration of integrated circuits functioning in the gigahertz frequency regime. However, the 2DLMs circuits reported to date can only function in a few megahertz or even lower frequency regime. The difficulties in integrating high quality dielectrics and conducting subtractive lithography on atomically thin materials have prevented achieving 2DLMs transistors with optimized device geometry and performance. It is a well recognized challenge to integrate dielectrics with graphene because of the intrinsic incompatibility of graphene with typical oxide dielectrics or their deposition approaches. Despite several attempts to date, the integration of high quality high-k dielectrics on TMDs such as MoS₂ remains elusive. Furthermore, another significant challenge to achieve high performance devices based on these atomically thin materials is their intrinsic incompatibility with the conventional subtractive lithography processes (e.g. various
plasma etching) that can severely damage the atomic structure and degrade the electronic properties.

Here we report the best performed MoS$_2$ transistors to date by using an additive lithography approach to integrate few-layer MoS$_2$ with transferred gate stacks$^9$. The transfer-gate strategy can allow a damage-free process to integrate MoS$_2$ with high quality dielectrics and self-aligned gate to achieve MoS$_2$ transistors with optimized device geometry and unprecedented performance, including a highest transconductance exceeding 60 S/m, an intrinsic gain over 30 and an on-off ratio exceeding $10^7$. On-chip microwave measurements demonstrate a highest intrinsic cut-off frequency $f_T$ of 42 GHz and a maximum oscillation frequency $f_{MAX}$ of 50 GHz, exceeding that of graphene transistors ($f_{MAX}$~44 GHz)$^{23}$. Furthermore, we have, for the first time, demonstrated the integration of multiple MoS$_2$ transistors on quartz and flexible substrates to form a logic inverter or radio frequency amplifier with voltage gain in the gigahertz regime.

**B. Fabrication of self-aligned MoS$_2$ transistors with transferred gate stacks**

Our devices are based on mechanically exfoliated few-layer MoS$_2$ flakes on Si/SiO$_2$ (300 nm) substrate, which were characterized by using optical microscope, atomic force microscopy (AFM), and Raman spectroscopy$^{2,3,28}$. The same process should be applicable to the large MoS$_2$ or other 2DLMs grown by chemical vapour deposition approach. The gate stacks – metal bars wrapped in dielectrics – were first patterned on a sacrificial substrate, and transferred onto the few-layer MoS$_2$$^9$. Following electron-beam lithography, a metallization (titanium (50nm)/gold (50nm)) process was used to define the external source, drain, and gate electrodes. A thin layer of Ni/Au (5nm/10 nm) was then deposited to form the self-aligned source and drain electrodes with minimized access resistance or parasitic capacitance$^{9,29}$ (Fig. 4-1a). Figure 4-1b shows a scanning electron microscopy (SEM) image of the top view of top-gated dual-channel self-aligned MoS$_2$
field effect transistors (FETs). The cross-sectional transmission electron microscope (TEM) image shows that the self-aligned source and drain electrodes are well separated and precisely positioned next to the gate spacer dielectrics (Fig. 4-1c). High resolution TEM image shows an 8-layer MoS$_2$ flake with clear interface between the transferred gate stack and MoS$_2$ surface (Fig. 4-1d). Since phonon scattering and roughness scattering can severely degrade the mobilities in atomically thin 2D materials, a high-$k$ dielectric with clean interface can screen the scatterings and enhance the mobility of MoS$_2$ devices$^3$.

The basic electronic properties of MoS$_2$ FETs were first probed using standard back-gate devices on Si/SiO$_2$ substrate (without top-gate). The transfer characteristics are determined by measuring the drain-source current $I_{ds}$ as a function of sweeping the back-gate voltage $V_{BG}$ at a fixed drain voltage $V_{ds}$ (Fig. 4-2a). The field-effect mobility of the device can be derived using $\mu = (dI_{ds}/dV_{BG})/(L/W \cdot C_0 \cdot V_{ds})$, where channel length $L = 1 \mu$m, channel width $W = 4 \mu$m, and back-gate capacitance $C_0 = \varepsilon_0 \varepsilon_r / d = 11.5$ nF/cm$^2$. A field-effect mobility of $\mu = 170$ cm$^2/(V\cdot s)$ at a drain voltage of $V_{ds} = 0.5$ V can be derived from the transconductance curve shown in right axis of Figure 2a. This value is comparable to the best mobility values in MoS$_2$ FETs reported to date$^{1,3,5}$.

In order to evaluate the layer thickness dependent electronic properties, we have measured more than 40 MoS$_2$ devices in back-gate configurations. In general, a thinner MoS$_2$ can usually offer a lower off-state current but a thicker one has a larger on-state current$^{10}$. The decreasing mobility with reducing thickness may be explained by using a generalized Coulomb scattering model$^{30}$. For even thicker MoS$_2$ (>20 nm), the back gate can hardly turn the channel on or off due to the screening effect of the bottom layers$^{31}$. A plot of the mobility versus the thickness shows that a MoS$_2$ with a thickness between 2-7 nm holds both the high mobility and high on-off ratio (Fig. 4-
Therefore, we have mainly focused on the MoS$_2$ flake with a thickness between 2-7 nm for the fabrication of high performance MoS$_2$ FETs in the following studies.

To probe the performance of limit of few-layer MoS$_2$ FETs, we have fabricated top-gated short channel devices with self-aligned source and drain electrodes (Ni/Au) to minimize the access resistance and parasitic capacitance. The output characteristics of a 100 nm channel length self-aligned MoS$_2$ transistor were measured at various gate voltages (Fig. 4-2c). The $I_{ds}$-$V_{ds}$ curve shows linear behaviour at low bias voltages, suggesting that the self-aligned Ni/Au thin film forms Ohmic contacts with MoS$_2$\textsuperscript{29}. Significantly, a clear current saturation is observed at high source drain bias, which is difficult to achieve at such short channel length in graphene based transistors. The drain-source conductance $g_{ds} = dI_{ds}/dV_{ds}$ is close to zero in this region of operation. Current saturation is a very important parameter for achieving maximum possible operating speeds\textsuperscript{32}.

The transfer characteristics ($I_{ds}$-$V_{TG}$ curves) of the same device were measured at different drain bias (Fig. 4-2d). Our study shows typical n-type FET characteristics with the threshold voltage located around -2V. Importantly, an on/off ratio exceeding $10^7$ is achieved in this device, sufficient for digital circuits\textsuperscript{33}, which cannot be achieved in graphene based devices. The subthreshold swing, defined as $S = (dV_{TG}/d(\log I_{ds}))$, $S = 83$ mV/dec can be extracted at $V_{ds} = 1$ V. In conventional MOSFETs, the subthreshold swing depends on the ratio of gate capacitance to the other parasitic capacitance such as interface trap state capacitance, and has a theoretical limit of 60 mV/dec at room temperature. A low subthreshold swing is generally desirable for low power low threshold operation. The subthreshold swing observed in our device is very close to the theoretical limit, which can be attributed to the highly quality MoS$_2$-dielectric interface obtained with our transfer-gate approach. The transconductance, defined $g_{m} = dI_{ds}/dV_{TG}$, can also be derived from $I_{ds}$-
$V_{TG}$ characteristics curves (Fig. 4-2e). A peak scaled transconductance of 60 $\mu$S/$\mu$m is obtained at $V_{ds} = 5$ V, which represents the highest transconductance value reported for MoS$_2$ FETs to date$^{34}$.

The ratio of transconductance to drain-source conductance defines the intrinsic gain ($A = g_m/g_{ds}$), which is an important figure-of-merit representing the highest achievable gain in a single transistor$^{18}$. With record high values of transconductance ($g_m = 60$ $\mu$S/$\mu$m for $V_{ds} = 5$ V) and clear current saturation ($g_{ds} < 2$ $\mu$S/$\mu$m) in these MoS$_2$ transistors, an intrinsic gain over 30 can be achieved. This is in stark contrast to graphene transistors with similar channel length in which intrinsic gain cannot be achieved due to little current saturation in output characteristics. Such high intrinsic gain observed in the MoS$_2$ transistors can address the critical limitation of graphene based transistors to open up exciting potential for both digital and analog applications with voltage gain.

We have also studied self-aligned MoS$_2$ transistors with variable channel lengths. The $I_{ds}$-$V_{TG}$ transfer curves were measured from more than 30 devices with self-aligned gate lengths ranging from ranging from 68 nm to 1 $\mu$m. The distributions of on-state current and transconductance with different channel lengths were extracted from the $I_{ds}$-$V_{TG}$ transfer curves. In general, both the on-state current and the transconductance increase with decreasing channel length (Fig. 4-2f), suggesting that the MoS$_2$ channel is dominating the charge transport and the self-aligned Ni/Au contacts form good contacts with MoS$_2$. Importantly, even the shortest channel device (68 nm) exhibited the current saturation with an intrinsic gain as large as 30, which can not be achieved in graphene devices with similar channel length.

**C. RF performance of self-aligned MoS$_2$ transistors with transferred gate stacks**

The above discussion clearly demonstrates that our self-aligned devices exhibit the best DC performance achieved in MoS$_2$ FETs to date. To further determine the cut-off frequency ($f_T$) of
these devices, we have conducted the on-chip microwave measurements with an Agilent 8361A network analyzer in the range of 50 MHz – 30 GHz. The MoS₂ transistors for radio frequency (RF) measurement were fabricated on highly resistive silicon substrate (>18,000 Ω·cm) in order to minimize the parasitic capacitance. To accurately assess the intrinsic $f_T$ value, careful de-embedding procedures were performed using the exact pad layout as “open”, “short” and "through" structures on the same chip. The de-embedded S parameters constitute a complete set of coefficients describing intrinsic input and output behaviour of MoS₂ transistors. The small signal current gain $|h_{21}|$ extracted from the measured S parameters exhibits a typical $1/f$ frequency dependence expected for an ideal FET (Fig. 4-3a). The linear fit yields cut-off frequencies $f_T = 13.5$ GHz, $26$ GHz and $42$ GHz for the transistors with channel lengths of $216$ nm (blue triangle in Fig. 3a), $116$ nm (red circle) and $68$ nm (black square), respectively. These values were further verified by using Gummel’s approach (Fig. 4-3a inset). We have also carefully analyzed the S parameters to derive the device component values (including gate-source capacitance, gate-drain capacitance and transconductance), which are consistent with those obtained from electrostatic simulations or DC measurements, demonstrating the validity of the RF measurements and the de-embedding procedures. To the best of our knowledge, the observed $f_T$ value of $42$ GHz is the highest cut-off frequency obtained in all TMD-based transistors reported to date.

To further investigate the reproducibility of our approach and examine the length-scaling relations, we systematically examined more than 20 MoS₂ transistors of variable channel lengths (Fig. 4-3b). In general, the peak cut-off frequencies follow $1/L$ dependence. It is valid for devices with short gate lengths, even at the scaling limit of $68$ nm. Although $1/L^2$ dependence was expected in theoretical model, a similar $1/L$ dependence is usually observed in short channel conventional Si and III–V FETs. This dependence is mainly due to the nearly-constant effective carrier velocity. 
obtained by reaching the saturation velocity of the material. The $1/L$ scaling trend observed in our devices indicates that the MoS$_2$ devices work in the saturation region of operation, which is very important for practical device operation. We believe the RF performance of our device can be further enhanced by further optimizing the contact resistance and/or improving gate coupling.

In addition to cut-off frequency ($f_T$), maximum oscillation frequency ($f_{\text{MAX}}$), defined as the frequency at which the power gain is equal to one, is another important figure-of-merit defining the RF performance of a transistor. The maximum available gain (MAG) was extracted from the measured S parameters (Fig. 4-3c). Importantly, $f_{\text{MAX}}$ of 16 GHz, 34 GHz and 50 GHz can be achieved in MoS$_2$ transistors with channel lengths of 216 nm (blue triangle) 116 nm (red circle) and 68 nm (black square), respectively. These values greatly exceed the best values reported for MoS$_2$ transistors to date ($f_T\sim$900 MHz, $f_{\text{MAX}} \sim$ 1 GHz). Further more, it is particularly important to note that the achieved $f_{\text{MAX}}$ of 50 GHz is also higher than the best value achieved in graphene transistors. To the best of our knowledge, this is the highest maximum oscillation frequency obtained in any 2DLMs to date. Similar to the case of $f_T$, $f_{\text{MAX}}$ possesses a monotonic behaviour when the channel length decreases (Fig. 4-3d). However, the trend of $f_{\text{MAX}}$ doesn’t follow the $1/L$ dependence. This is the result of competing contributions from $f_T$, gate resistance, and output conductance $g_{\text{ds}}$ as the gate length decreases. To further improve the power gain, one needs to improve the cut-off frequency through using MoS$_2$ devices with higher mobility, or decrease the gate resistance by using T-gate, or improve the saturation behaviour of the MoS$_2$ devices through gate dielectric down-scaling.

**D. Circuits based on MoS$_2$ transistors working in the gigahertz regime**

With excellent on-off current ratio, intrinsic gain, intrinsic cut-off frequency and power gain performance, the MoS$_2$ transistors can enable exciting opportunities for both the digital and analog
electronics. To this end, we have fabricated an inverter circuit by connecting an enhancement-mode (E-mode) MoS$_2$ transistor with a depletion-mode (D-mode) resistor (formed by connecting the gate of a depletion-mode transistor directly to source electrode) (Fig. 4-4a). To minimize the parasitic capacitance for high frequency measurement, the inverter circuit was fabricated on quartz substrate. An inverter circuit is a basic logic element that outputs a voltage representing the opposite logic-level to its input. The quality of a logic inverter is often evaluated using its voltage transfer curve (Fig. 4-4b), which is a plot of input voltage versus output voltage. When the input voltage is relatively high (logic state 1), the E-mode MoS$_2$ FET is much more conductive than the D-mode FET, setting the output voltage close to 0 V (logic state 0). When the input voltage is relatively low (logic state 0), the D-mode MoS$_2$ FET is much more conductive than E-mode FET and the output is set close to 5 V (logic state 1). The slope of the transition region in the middle defines voltage gain. Importantly, a voltage gain $> 10$ is achieved in our MoS$_2$ based inverter circuit (Fig. 4-4b). The achievement of such high gain demonstrates that self-aligned MoS$_2$ devices can be used for the fabrication of integrated circuits and for high performance logic operations at room temperature.

Although MoS$_2$ transistor based circuits have been demonstrated previously, these circuits typically operate in the relatively low frequency regime up to a few megahertz$^{10,12,13,38}$. With the maximum oscillation frequency up to 50 GHz, our self-aligned MoS$_2$ transistors can be readily used to construct RF circuits up to the gigahertz frequency regime. For example, with input signal of 200 MHz square wave applied to the input electrode of our MoS$_2$ inverter, an inverted signal with a relative voltage gain of two can be obtained at an operating frequency of 200 MHz without any noticeable delay (Fig. 4-4c). Since the resistance of MoS$_2$ inverter can be tuned by the DC gate voltage offset applied on the input electrode, the output gain and frequency response of the MoS$_2$
inverter can be tuned by changing the DC voltage offset applied on the input electrode. It is important to note that our inverter performance does not exclude any parasitic capacitances (such as $C_{gs}$, $C_{gd}$ etc.) or series inductances, which highly depend on the exact circuit design. A better circuit design or a thinner dielectric layer or a larger bias can further improve the high frequency performance of MoS$_2$ based inverter.

With much higher voltage gain than graphene transistors, the MoS$_2$ transistors can also be used to construct RF amplifiers (Fig. 4-4d). The amplifier is obtained by integrating two transistors in series, where the upper one acts as a “switch” and the lower one acts as an active “load”. The gate of “switch” transistor acts as input, whereas the gate of the “load” transistor is connected to the central lead and acts as the output. In order to maximize the performance of the amplifier, the power supply of the amplifier is set at 6V. A small sinusoidal signal $V_{\text{IN}}$ is superimposed on the DC bias $V_g$ via a bias-T. The output signal $V_{\text{OUT}}$ is connected to an oscilloscope via a DC blocker. The circuit is first biased at a certain DC bias and gate voltage to establish a desired current in the circuit. When a small sinusoidal signal is applied on the input electrode, the circuit can work as a linear amplifier. As shown in Figure 4-4e, a 100 MHz sinusoidal wave with amplitude of 100 mV peak-to-peak voltage was applied on the input electrode. A relative voltage gain of 2 can be observed in the output signal, which is very important for practical application. Furthermore, the amplifier exhibited a larger than unit gain (1.07) with an input of sinusoidal wave with a frequency of 1 GHz (Fig. 4-4f), demonstrating our amplifier can work in the gigahertz regime with voltage gain, which is advantageous over graphene transistors with which the voltage gain is lacking due to the lack of current saturation. We have performed measurements from 60 MHz to 1.5 GHz and determined the relative voltage gain of our amplifier versus the frequency (Fig. 4-4g). It is clear
that our amplifier preserve the relative voltage gain equal to 2 up to 500 MHz, and retains a relative voltage gain >1 at 1 GHz.

The propagation delay of MoS$_2$ amplifier can be probed by applying a square wave generated by an arbitrary waveform generator on the input electrode and measuring the output voltage response by using an oscilloscope. When an input signal with a rise time of 210 ps was applied on the input electrode of our MoS$_2$ amplifier, an output voltage with a rise time of 580 ps was captured by the oscilloscope (Fig. 4-4h). Considering the delay of input signal, a propagation delay of 370 ps is observed in our MoS$_2$ amplifier. A square wave is a non-sinusoidal periodic waveform, which can be represented as an infinite summation of sinusoidal waves. By considering the gain for each component of sinusoidal wave, we can simulate the output signal of the propagation delay measurement (detailed calculation in Supplementary equation S1). Importantly, the simulated output signal matches well with the experimental result, indicating the proper functionality of our MoS$_2$ amplifier in gigahertz regime (Fig. 4-4i).

2DLMs are promising candidates for both flexible and stretchable electronics applications, such as low-power, high-frequency electronics, optoelectronics, and integrated systems$^{40,41}$. Graphene has been widely speculated for high performance flexible electronics due to its extremely high carrier mobility and excellent mechanical properties$^{42}$, but is limited by its semimetal nature and the lack of intrinsic voltage gain. With excellent semiconducting characteristics and few-atomic thickness, MoS$_2$ is considered an ideal material for high speed flexible electronics. Although MoS$_2$ has been explored for flexible transistors, but typically in DC regime to date$^{41,43,44}$. Importantly, our fabrication approach can be readily applied onto flexible substrate to enable high performance MoS$_2$ transistors for low-power flexible electronics (Fig. 4-5a). The self-aligned MoS$_2$ transistors on flexible substrate exhibit a similar performance with a
highest current density of 48 \( \mu A/\mu m \) achieved in a 100 nm channel length MoS\(_2\) transistor on flexible substrate (Fig. 4-5b), greatly exceeding the recent published results\(^{11}\). An inverter circuit made from MoS\(_2\) transistor exhibit a very sharp transition with a voltage gain \( \sim 9 \) (Fig. 4-5c). Additionally, we have further tested the intrinsic RF performance by measuring the intrinsic cut-off frequency \( (f_t) \) and maximum oscillation frequency \( (f_{\text{MAX}}) \) of the MoS\(_2\) FETs on flexible substrate. An intrinsic cut-off frequency \( (f_t) \) of 13.5 GHz and maximum oscillation frequency \( (f_{\text{MAX}}) \) of 10.5 GHz are achieved in a 68 nm MoS\(_2\) FET (Fig. 4-5d). The maximum oscillation frequency obtained here also exceeds the best results achieved in graphene flexible transistors \( (f_{\text{MAX}} \sim 3.7 \text{ GHz})^{45,46} \). These studies demonstrate that MoS\(_2\) may be used as an ideal material for flexible electronics that requires both low power and high speed characteristics.

Taking a step forward, we have also constructed an RF amplifier on flexible substrate based on self-aligned MoS\(_2\) FETs with a channel length of 100 nm and determined its output characteristics. A larger than unit relative voltage gain was observed in the output signal for a 300 MHz sinusoidal input signal (Fig. 4-5e). Our study represents the first demonstration of MoS\(_2\) circuit on flexible substrate, and as far as we know, there is yet no report of flexible graphene RF circuit with voltage gain above 1 MHz to date. In order to measure the propagation delay of MoS\(_2\) based circuits on flexible substrate, a MoS\(_2\) based inverter was fabricated with MoS\(_2\) FETs with a channel length of 100 nm. The propagation delay of MoS\(_2\) amplifier on flexible substrate was measured by applying an input signal with a rise time of 0.2 ns on the input electrode of MoS\(_2\) based inverter (Fig. 4-5f, black curve). An inverted output voltage with a rise time of 2.2 ns was captured by the oscilloscope (Fig. 4-5f, red curve). Together, a propagation delay of 2 ns is achieved in our MoS\(_2\) amplifier on flexible substrate, demonstrating exciting potential of MoS\(_2\) transistors for flexible RF applications.
E. Summary

In summary, we have demonstrated the best performed few-layer MoS$_2$ transistors to date with on-off ratio over $10^7$, intrinsic gain up to 30, intrinsic cut-off frequency up to 42 GHz and maximum power gain performance up to 50 GHz on Si/SiO$_2$, quartz and flexible substrate. Importantly, with an intrinsic band gap, the MoS$_2$ based transistors can offer several advantages compared to the graphene transistors, including large on/off ratio, excellent current saturation, large intrinsic gain, and greatly better power gain performance. Exploiting these unique advantages, we have demonstrated that the few-layer MoS$_2$ transistors can be used to construct functional circuits, including logic inverter and RF amplifier, operating in the gigahertz regime with voltage gain, which is difficult to achieve in graphene transistor based RF circuit. Our study represents a first demonstration of MoS$_2$ based gigahertz circuits with current saturation and voltage gain, and establishes an important milestone in applying 2DLMs for high-performance electronics, particularly flexible electronics.
References


32 Tsividis, Y.; McAndrew, C. *Operation and modeling of the MOS transistor*. (Oxford University Press, 2010).

33 The international technology roadmap for semiconductors (Semiconductor Industry Association, 2011).


38 Wang, H. *et al.* Large-scale 2D electronics based on single-layer MoS$_2$ grown by chemical vapor deposition. *2012 IEDM Tech Dig.* S4.6 (2012).
39 Jan, C.-H. et al. RF CMOS technology scaling in high-\(k\)/metal gate era for RF SoC (system-on-chip) applications. 2010 IEDM Tech Dig. 27.2 (2010).


Figures & Legends

Figure 4-1 | Schematic illustration and characterization of the MoS$_2$ FETs with transferred gate stacks. a, A schematic illustration of the MoS$_2$ FETs with transferred gate stacks, and the inset show schematic illustration the cross-sectional view of the self-aligned device. b, The SEM image of MoS$_2$ FETs with transferred gate stacks. The scale bar is 5 μm. c, The Cross-sectional TEM image of the overall device layout. The scale bar is 50 nm. d, HRTEM image of the interface between MoS$_2$ and transferred gate stack. The scale bar is 3 nm.
Figure 4-2 | Room-temperature DC characterization of the self-aligned MoS$_2$ FETs with transferred gate stacks. **a,** The $I_{ds}$-$V_{BG}$ transfer characteristics and the corresponding transconductance of the device at 0.5 V bias voltage for the 1 μm channel length back-gated MoS$_2$ transistor. **b,** The distribution of MoS$_2$ transistor mobility and on-off ratio in back-gate configuration versus thickness of MoS$_2$. **c,** $I_{ds}$-$V_{ds}$ output characteristics at various gate voltages ($V_{TG}$ from -2 V to 4 V) for a 100 nm channel length top-gated MoS$_2$ transistor with self-aligned source and drain electrodes. **d,** The $I_{ds}$-$V_{TG}$ transfer characteristics at different bias voltage for a 100 nm channel length top-gate configuration MoS$_2$ transistor with self-aligned source and drain electrodes ($V_{ds}$= 1, 2, 3, 4 and 5 V). **e,** The corresponding transconductance of the MoS$_2$ transistor shown in Fig. 3e at different bias voltage. **f,** The distribution of on-state current and transconductance versus channel length in top-gated MoS$_2$ transistors with self-aligned source and drain.
Figure 4-3 | Radio frequency performance of the self-aligned MoS$_2$ FETs with transferred gate stacks. 

a, Small-signal current gain $|h_{21}|$ versus frequency for three devices with a channel length of 216 nm (blue), 116 nm (red), and 68 nm (black) at room temperature. The cut-off frequencies are 13.5 GHz, 26 GHz and 42 GHz respectively at a DC bias of 5 V. 

b, Peak $f_T$ as a function of gate length from over 20 MoS$_2$ FETs. 

c, Maximum Available Gain (MAG) versus frequency for three devices shown in Figure 3a with a channel length of 216 nm (blue), 116 nm (red), and 68 nm (black) at room temperature. The Maximum Available Gain are 16 GHz, 34 GHz and 50 GHz respectively at a DC bias of 5 V. 

d, Peak $f_{MAX}$ as a function of gate length from over 20 MoS$_2$ FETs.
**Figure 4-4** Demonstration of an integrated logic inverter and signal amplifier based on few-layer MoS$_2$ FETs with transferred gate stacks on quartz substrate. 

**a,** Schematic illustration an integrated logic inverter made two MoS$_2$ transistors. 

**b,** Output voltage of the integrated logic MoS$_2$ inverter as a function of the input voltage (left axis), and the gain of the integrated logic MoS$_2$ inverter (right axis), highlighting a voltage gain $>10$. The inset shows an optical image of an integrated logic inverter on quartz substrate. 

**c,** Input (black) and output (red) signal of the few-layer MoS$_2$ inverter. The input signal is a 200 MHz square wave signal with amplitude of 100 mV and DC gate bias. The output signal is shifted in phase for 180° with a gain of 2 over the input signal. 

**d,** Schematic of an integrated RF amplifier made by integrating two MoS$_2$ FETs. 

**e,** 100 MHz sinusoidal input signal (black) coupled with DC gate bias and the output signal with a voltage gain of 2. 

**f,** 1 GHz sinusoidal input signal (black) coupled with DC gate bias and the output signal with a voltage gain of 1.07. 

**g,** The frequency dependence of the small signal voltage gain in few-layer MoS$_2$ amplifier. 

**h,** Propagation delay of few-layer MoS$_2$ amplifier, with the input signal (black) showing a rise time of 210 ps and output signal (red) showing a rise time of 580 ps. 

**i,** Simulation of the propagation delay based on the frequency dependent gain of few-layer MoS$_2$ amplifier.
Figure 4-5] Demonstration of an integrated logic inverter and signal amplifier based on few-layer MoS₂ FETs with transferred gate stacks on flexible substrate. **a**, A photo of MoS₂ circuits on flexible substrate. **b**, The $I_{ds}$-$V_{bg}$ transfer characteristics of the device at 2 V bias voltage for the 100 nm channel length top-gated MoS₂ transistor on flexible substrate. **c**, Output voltage of the integrated logic MoS₂ inverter as a function of the input voltage (left axis), and the gain of the integrated logic MoS₂ inverter (right axis). **d**, Small-signal current gain ($|h_{21}|$), Mason's Unilateral Power Gain (U), and Maximum Available Gain (MAG) versus frequency for a 68 nm channel length MoS₂ transistor on flexible substrate at a DC bias of 8 V. **e**, A 300 MHz sinusoidal input signal (black) coupled with DC gate bias and the output signal with a voltage gain of 1.05. **f**, Propagation delay of few-layer MoS₂ inverter. The input signal (black) has a rise time of 0.2 ns and the output signal (red) has a rise time of 2.2 ns.
Figure 4-6 | Optical microscopy, atomic force microscopy, and Raman microscopy characterization of MoS$_2$ on 300 nm SiO$_2$ substrate. a, Optical image of MoS$_2$ on 300 nm SiO$_2$ substrate. Scale bar is 10 μm. b, AFM image of the same piece of MoS$_2$ on 300 nm SiO$_2$ substrate. c, Cross-sectional plot of height along the line in Fig. S1b. The curve shows that the thickness of MoS$_2$ flake is ~ 2.0 nm. d, Layer dependent Raman spectra of MoS$_2$ flakes on 300 nm SiO$_2$ substrate. The space between $A_{1g}$ mode and $E_{2g}$ mode is used to determine the thickness of MoS$_2$ flake.

Figure 4-7 | The $I_{ds}$-$V_{BG}$ transfer characteristic curves for back-gated MoS$_2$ transistors with different thickness. The channel length of these transistors is 1μm. The bias applied on the transistors is $V_{ds}$=0.5 V. In general, a thinner MoS$_2$ can usually offer a lower off-state current but a thicker one has larger on-state current. But the off-state current also increases with increasing
thickness. Few-layer MoS$_2$ (~ 2-7nm) can offer the best balance between the off-state and on-state current.

![Electric potential distribution](image)

Figure 4-8| Finite element simulation of electrostatic capacitance between the gate stack and MoS$_2$ in a 68 nm top-gate device. The simulated electrostatic capacitance normalized by MoS$_2$ channel width is 0.271 fF/ m, which is consistent with the component parameters calculated in Supplemental Table 4-1.

<table>
<thead>
<tr>
<th>Width (μm)</th>
<th>Length (nm)</th>
<th>g$_m$ (mS)</th>
<th>C$_{gs}$ (fF)</th>
<th>C$_{gd}$ (fF)</th>
<th>R$_s$ (Ω)</th>
<th>R$_d$ (Ω)</th>
<th>Projected f$_T$ (GHz)</th>
<th>f$_T$ (GHz)</th>
<th>f$_{MAX}$ (GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>68</td>
<td>1.46</td>
<td>4.9</td>
<td>0.43</td>
<td>24</td>
<td>21</td>
<td>44</td>
<td>42</td>
<td>50</td>
</tr>
</tbody>
</table>

Table 4-1| The values of component parameters of a 68 nm channel length MoS$_2$ device in RF measurement.

Equation 4-1| The simulation of the output signal propagation delay in few-layer MoS$_2$ based amplifier.

The input square wave can be represented as an infinite summation of sinusoidal waves as shown below:

\[ x_{\text{square}}(t) = \frac{4}{\pi} \sum_{k=1,3,5,\ldots}^{\infty} \frac{\sin(2\pi ft \times (2k-1))}{(2k-1)} = \frac{4}{\pi} (\sin(2\pi ft) + \frac{1}{3} \sin(6\pi ft) + \frac{1}{5} \sin(10\pi ft) + \ldots) \]

For each component of sinusoidal wave, the corresponding output signal can be expressed by multiplying the input sinusoidal wave by amplitude of the amplifier measured at certain frequency.
\[ V_{OUT}(f) = A(f) \times V_{IN}(f) \]

The simulated output signal can be represented as summation of each corresponding output sinusoidal wave:

\[ V_{OUT} = \frac{4}{\pi} (\sin(2\pi f t) \times A_1(f) + \frac{1}{3} \sin(6\pi f t) \times A_3(3f) + \frac{1}{5} \sin(10\pi f t) \times A_5(5f) + \cdots) \]
A. Introduction to van der Waals heterojunction

Two-dimensional layered materials, such as graphene, MoS$_2$, and WSe$_2$, are emerging as an exciting material system for future optoelectronics, including photodetectors$^{1-7}$, ultrafast lasers$^8$, polarizers$^9$, touch panels$^{10}$ and optical modulators$^{11}$ due to their atomically thin geometry and unique optical, electronic properties$^{12-23}$. In this regard, the monolayer transition metal dichalcogenides (ML-TMDs) is particularly interesting due to their direct energy bandgap and the non-centrosymmetric lattice structure$^{12-13}$. Electroluminescence (EL) from ML-MoS$_2$ has been reported in a metal-MoS$_2$ Schottky junction through a hot carrier process$^{24}$. It has also been demonstrated that ML-TMDs can be electrostatically doped by applying different bias to capacitively coupled gate electrodes to form a planar p-n junction$^{25-27}$, but typically with relative low optoelectronic efficiency (~0.1–1% external quantum efficiency (EQE) for photon to electron conversion).

The atomically thin geometry of these 2D materials can allow band structure modulation in a vertically stacked heterostructures to form atomically sharp junctions$^{28}$. For example, this strategy allows gapless graphene to be used in field-effect tunnelling devices$^{28,29}$, barristors$^{30}$, inverters$^{31}$, and photodetectors$^7$ while staked with other 2D materials in the vertical direction. Although the nearly perfect 2D structure and low density of states in graphene provide advantages in some heterostructure devices, its gapless nature prevents the formation of a large potential barrier for charge separation and current rectification. The vertical heterojunction p-n diode formed between one TMD material and a bulk material has recently been reported, but usually with no EL$^{32,33}$ or very weak EL$^{34}$. As far as we know, there is yet no report of vertically stacked p-n diodes formed between two different atomic layered materials to date.
Here we report an atomically thin p-n diode based on a heterojunction between synthetic $p$-type ML-WSe$_2$ and exfoliated $n$-type MoS$_2$ flake. A novel structure of vertical heterojunction p-n diode is achieved using two atomically thin TMD materials for the first time. The atomically thin p-n diode exhibit well defined current rectification behaviour and can enable efficient photocurrent generation with an EQE up to 12%. Unlike the planar structures where the active area is confined to the lateral interface region, photocurrent mapping of our device demonstrates that the p-n junction is created for the entire overlapping area with excellent diode characteristics. Furthermore, prominent EL is observed under forward bias. A systematic investigation of the EL spectra also reveals important insights about electron-orbital interaction in TMD based materials.

B. Fabrication and characterization of the heterojunction p-n diode.

B1. Fabrication of the heterojunction p-n diode

The vertical heterojunction p-n diode is formed between synthetic $p$-type ML-WSe$_2$ and exfoliated $n$-type MoS$_2$ flake (Fig. 5-1a,b). Triangular domains of ML-WSe$_2$ was first synthesized on 300 nm Si/SiO$_2$ substrate typically with a bilayer (BL) region at the centre (see Fig. 5-2a), which were characterized by using optical microscope, atomic force microscopy (AFM), and Raman spectroscopy. Mechanically exfoliated MoS$_2$ flakes were then transferred onto synthetic WSe$_2$ domains to form vertically stacked heterojunctions. Electron-beam lithography and electron beam evaporation was used to define the contact electrodes. A thin Ni/Au film (5nm/50 nm) and Au film (50nm) were used as the electrode for MoS$_2$ flake$^{35}$ and WSe$_2$ domain to form Ohmic contacts with minimized contact resistance and potential barrier (Fig. 5-1b). Figure 5-1c shows the ideal band diagrams of the heterojunction p-n diode at zero bias. The built-in potential and applied voltage are mainly supported by a depletion layer with abrupt atomic boundaries, and outside the boundaries the semiconductor is assumed to be neutral. Under zero bias, $p$-type WSe$_2$ and $n$-type
MoS₂ form a depletion region at the interface preventing both electron and hole injection into the other side of the heterojunction.

**B2. Structural characterization of the heterojunction p-n diode.**

Figure 5-2a shows an optical microscopy image of a synthetic WSe₂ domain on 300nm Si/SiO₂ substrate. A triangular shaped BL-WSe₂ was typically observed at the center of the triangular ML-WSe₂ domain, indicating the nearly perfect lattice structure of our synthetic WSe₂. Figure 5-2b shows a top-view scanning electron microscopy (SEM) image of the vertical heterojunction. The MoS₂, WSe₂ layers and the contact electrodes are labelled with different artificial colour to highlight the device structure. We also carried out photoluminescence (PL) mapping to further illustrating the stacking structure of WSe₂/MoS₂ heterojunction (Fig. 5-2c). The PL mapping show distinct PL from WSe₂ (red region in Fig. 5-2c) and that from MoS₂ (green region in Fig. 5-2c), consistent with structure layout observed in the SEM image (Fig. 5-2b). The uniform PL from WSe₂ and MoS₂ also indicates the excellent material quality of our TMD based materials. The PL spectra of WSe₂ show a strong layer-number dependence (Fig. 5-2d), with the PL intensity in ML-WSe₂ at least more than 10 times stronger than that in BL-WSe₂. The PL spectrum in ML-WSe₂ shows a peak at 785 nm, corresponding to the “A” exciton peak. The PL in BL-WSe₂ also exhibits the A exciton peak with an additional broad peak at ~877 nm, which is attributed to indirect band gap emission involving a conduction band minimum at a midpoint between K and Γ points and valence band maximum at the Γ point (typically label as “I” peak). These PL studies are consistent with previous experimental studies and theoretical calculations, indicating the good crystalline quality of our synthetic WSe₂. The PL spectrum from MoS₂ flake shows a peak at 677 nm, corresponding to “A” exciton peak in MoS₂. It is also important to note that the “B”
exciton peak can be observed in both WSe₂ (605 nm) and MoS₂ (620 nm), but with the intensity 2-3 order magnitude lower than the A exciton peak (Fig. 5-2d inset).

We have further characterized the stacking structure of the hetero-junction using cross-sectional transmission electron microscope (TEM) studies. The high resolution TEM image clearly shows the WSe₂/MoS₂ heterojunction with a 13-layer MoS₂ flake on top of BL-WSe₂ (Fig. 5-2e). Energy dispersive X-ray spectroscopy (EDS) was further used to analyze the elemental distribution across the heterojunction interface. An EDX elemental line scan in vertical direction shows that W-L characteristic peaks were narrowly distributed in a specific region of our heterojunction (Fig. 5-2f). A careful analysis shows that the W-L distribution profile can be fitted by two Lorentz peaks with a peak separation of 0.7 nm, consistent with the observation of BL-WSe₂ in Fig. 5-2e. Together, these structural and PL characterisations demonstrates that a vertically stacked heterojunctions are formed by vertically stacking atomically thin WSe₂ and MoS₂.

**B3. Electrical characterization of the heterojunction p-n diode.**

Before testing the transport characteristics of the heterojunction p-n diodes, we have first characterized the electrical transport properties of MoS₂ and WSe₂ to ensure Ohmic contacts were achieved. To this end, the MoS₂ and WSe₂ field effect transistors (FETs) were fabricated on Si/SiO₂ substrate, with Ni/Au thin film as the source-drain contacts for MoS₂, and Au thin film as the contacts for WSe₂, and the silicon substrate as a back gate electrodes. Figure 5-3a and b show the $I_{ds}-V_{ds}$ characteristics at varying back gate voltage for MoS₂ and WSe₂, respectively. Importantly, a linear $I_{ds}-V_{ds}$ relationship is clearly observed for both MoS₂ and WSe₂, indicating Ohmic contacts are achieved for both materials. The formation of Ohmic contacts for both MoS₂ and WSe₂ is very important, since the Schottky barrier at the contact area may severely affect the electronic and optoelectronic characteristics of our vertical heterojunction and could induce
photocurrent generation or EL at the contact region. Furthermore, $I_{ds}-V_{ds}$ plot at varying back gate voltage show that the current increases with increasing positive gate voltage for MoS$_2$, indicating an n-type semiconductor behaviour. On the contrary, the current increases with decreasing negative gate voltage for WSe$_2$, demonstrating the p-type characteristics of WSe$_2$.

With Ohmic contacts formed for both MoS$_2$ and WSe$_2$, we continue to probe the electrical transport properties of the heterojunction p-n diode. Importantly, a clear current rectification behaviour is observed in $(I_{ds}-V_{ds})$ plots for the WSe$_2$/MoS$_2$ heterojunction (Fig 5-3c), with current only being able to pass through the device when the p-type WSe$_2$ is positively biased. The observation of current rectification clearly demonstrates a p-n diode is formed within the atomically thin WSe$_2$/MoS$_2$ heterojunction. The ultrathin nature of the heterojunction allows gate tunability of the diode characteristics. The diode output characteristic $(I_{ds}-V_{ds})$ under different back gate voltage show that the output current decreases with increasing positive gate voltage, suggesting that the p-type WSe$_2$ is limiting the charge transport when the diode is turned on.

The $I_{ds}-V_{ds}$ output characteristics of the vertical heterojunction under forward bias can be viewed as a vertical heterojunction p-n diode in series with an additional p-type FET due to the side contact on WSe$_2$ electrode. In general, the heterojunction p-n diode resistance decrease exponentially with increasing bias voltage, and the series p-FET resistance is nearly constant with bias voltage. Therefore, the heterojunction resistance is dominated by p-n diode at low bias and dominated by the p-type WSe$_2$ FET under high forward bias. We have also fitted the diode characteristics and calculated the ideality factor of our heterojunction device based on the model of a p-n diode with a series resistor (Fig. 5-3d). Importantly, an ideality factor of n=1.2 was derived with a series resistance of 80 MΩ, at zero gate voltage, and an ideality factor of n=1.3 was derived with a series resistance of 33 MΩ, at -20V gate voltage. The achievement of ideality factor close to
1 indicates the excellent diode behaviour of our atomically sharp heterojunction p-n diode. The decrease of the series resistance with increasing negative gate voltage is also consistent with our model that the p-type WSe₂ is the limiting series resistor at high forward bias.

C. Photocurrent and electroluminescence generation from WSe₂/MoS₂ heterojunction

C1. Photocurrent generation from WSe₂/MoS₂ heterojunction

The electrical measurements indicate the excellent diode behaviour in the atomically thin vertical heterojunction. To further characterize the diode characteristics in our vertical heterojunction, the photocurrent mapping was carried out at zero bias under a confocal microscope. Figure 5-4a shows an optical microscope image of the WSe₂/MoS₂ heterojunction depicting the relative position between WSe₂, MoS₂ and the electrodes. The corresponding photocurrent mapping at zero bias with a 514 nm laser excitation (5 µW) is shown in Figure 5-4b, with the ML-WSe₂ region outlined by purple square dotted line, few-layer MoS₂ outlined by blue rounded dotted line, and the electrodes outlined by golden solid lines. The photocurrent mapping show clear photoresponse from the entire overlapping region, indicating the formation of a broad area p-n junction across the entire overlapping area. It is also interesting to note that the photocurrent in ML-WSe₂/MoS₂ region is much stronger than that in BL-WSe₂/MoS₂ region, suggesting that the direct band gap plays an important role in the photocurrent generation process. A detailed understanding of the different response of ML vs. BL-WSe₂/MoS₂ will be an interesting topic for future studies. No measurable photocurrent was observed from the non-overlapping regions (only WSe₂ or MoS₂) or the electrical contacts, which is expected for zero bias photocurrent since the photogenerated carries in the regions outside p-n junction cannot be effectively separated and extracted.
The output characteristics ($I_{ds}$-$V_{ds}$) of the vertical heterojunction with and without laser illumination (514 nm, 5 µW) show clear photovoltaic power generation with an open-circuit voltage of $\sim$0.27 V and a short-circuit current of $\sim$ 0.22 µA (Fig. 5-4c). In general, the photoresponse exhibits a rapid temporal response beyond our experimental time resolution of 100 µs (Fig. 5-4c inset), demonstrating that the photoresponse is originated from photocarrier generation rather than any other extrinsic effects. Based on the photocurrent response and input laser power, we can determine the external quantum efficiency (EQE) of the photon to electron conversion. The EQE ($\eta$) is defined as the ratio of the number of carriers collected by electrodes to the number of the incident photon, or $\eta = (I_{ph}/q)/(P/h\nu)*100\%$ where $I_{ph}$ is the photocurrent, $h$ is Planck’s constant, $\nu$ is the frequency of light, $q$ is the electron charge and $P$ is the incident light power. Our study showed that the EQE in our vertical heterojunction can reach 11% under an excitation power of 5 µW 514 nm laser. Furthermore, it is found that the EQE increases with decreasing excitation laser power and decreases with increasing excitation power (Fig. 5-4d), with a maximum EQE of 12% observed under an excitation power of 0.5 µW. The decreasing EQE with increasing excitation power could be attributed partly to absorption saturation in WSe$_2$ and partly to the screening of a built-in electric field by the excited holes in the valence band of WSe$_2$.

The power dependent EQE of the same device under 633 nm excitation show a similar trend but with slightly lower values than those under 514 nm excitation, which may be attributed to the spectral dependent optical absorption coefficient. It is important to note that the EQE observed in the vertical WSe$_2$/MoS$_2$ heterostructure devices is much higher than those in lateral electrostatically doped WSe$_2$ p-n homojunctions (0.1-1%)$^{26-27}$, which may be partly attributed to more efficient charge separation resulting from an atomically sharp vertical p-n junction. In
contrast, the electrostatic doping would typically exhibit a spatial doping gradient, and is difficult to achieve atomically sharp junctions.

C2. Electroluminescence generation from WSe$_2$/MoS$_2$ heterojunction

The above electrical transport and photocurrent studies demonstrate excellent p-n diode characteristics in the atomically sharp WSe$_2$/MoS$_2$ heterojunction. Since p-n diode represents the basic device element for a light-emitting diode, we have further investigated the electroluminescence from these heterojunction p-n diodes. Figure 5-5a shows an EL image acquired under a forward bias of 3V and a forward current of ~100 μA. The shapes of WSe$_2$, MoS$_2$ and gold electrodes were outlined in the same way as before to identify the position of the EL. In contrast to the photocurrent generation from the entire overlapping area, it is important to note that the EL is localized at the overlapping area in close proximity to the electrodes. This can be explained by the electric field distribution in the heterojunction under different bias. For photocurrent mapping at zero bias (or a small bias less than the turn on voltage), the p-n diode junction resistance dominate the entire device, and therefore photocurrent can be seen from the entire overlapping area where there is a p-n junction. For EL studies at much higher forward bias exceeding the p-n diode turn-on voltage, the resistance of the ML-WSe$_2$ becomes an increasingly important component of the total resistance. Therefore, the most voltage drop occurs across the heterojunction edge near the electrodes due to the large series resistance of the ML-WSe$_2$. This is also consistent with the result reported for MoS$_2$/Si heterojunctions$^{34}$.

Figures 5-5b and c show the EL spectra of a ML- and a BL-WSe$_2$/MoS$_2$ heterojunction with increasing injection current. The plot of the overall EL intensity as a function of injection current shows an apparent threshold (Fig. 5-5d), with little EL below the threshold, and linear increase above threshold. The threshold current may be explained by the band alignment of the
heterojunction under different bias voltages (Fig. 5-5e,f). In general, due to different band gap and band alignment among the conduction band and valence band edge, the barrier for hole transport across the junction is smaller than that for the electron. With increasing forward bias (below a certain threshold), the holes from WSe\(_2\) are first injected into \(n\)-type MoS\(_2\) region, while few electrons can overcome the barrier to reach WSe\(_2\) (Fig. 5-5e). Due to the nature of indirect band gap in few-layer MoS\(_2\), the yield of radiative recombination is relatively low at this point. As a result, the EL intensity is very low when the hole injection dominates the charge transfer across the heterojunction. With further increasing bias across the heterojunction (above electron injection threshold), the conduction band of MoS\(_2\) is shifted upward, both electrons and holes can go cross the heterojunction and are injected into \(p\)-type and \(n\)-type region respectively (Fig. 5-5f). At this point, the radiative recombination in WSe\(_2\) dominates the EL with its intensity increasing linearly with the injection current. It is noted that the EL intensity observed in ML-WSe\(_2\)/MoS\(_2\) heterojunction is much stronger than that in BL-WSe\(_2\)/MoS\(_2\) heterojunction due to the higher radiative recombination rate in direct band gap ML-WSe\(_2\) vs indirect bandgap BL-WSe\(_2\).

The EL spectra show rich spectral features and can be well fitted using multiple Guassian functions with five main peaks, which can be assigned as excitonic peaks A (~792 nm) and B (~626 nm), hot electron luminescence (HEL) peaks A’ (~546 nm) and B’ (~483 nm) and an indirect band gap emission peak I (~880 nm). The A exciton peak dominates the spectra of the EL in ML-WSe\(_2\) (Fig. 5-5a) while the indirect band gap emission I is significant in BL-WSe\(_2\) (Fig. 5-5b). Strikingly, the EL spectra show prominent B exciton peak and HEL A’, B’ peaks in both ML- and BL-WSe\(_2\), which are usually 100-1000 times weaker than the A exciton peak in the PL measurements\(^{36}\) and have not been reported in EL previously. In contrast, the EL spectra of the vertical heterojunction show that the intensities of these HEL peaks are only 3-10 times weaker
than the A exciton peak, suggesting a relative enhancement of the HEL by about two orders of magnitude in EL, which may be attributed to the electric field induced carrier redistribution\textsuperscript{38}. The origin of the HEL peaks in TMD materials remains subject of debate and is difficult to probe due their low emission probability\textsuperscript{36, 39-41}. The HEL peaks A’ and B’ are generally believed to arise from the splitting of the ground and excited states of A and B transitions due to the electron-orbital interaction via either inter- or intralayer perturbation or both\textsuperscript{39,40}. However, there is no yet clear evidence to prove which perturbation dominates the electron-orbital interaction. The emergence of intense HEL emission in our ML-WSe\textsubscript{2}/MoS\textsubscript{2} and BL-WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction can offer a new platform to probe the origin of HEL peaks and the nature of electron-orbital interaction in TMDs. The presence of HEL peaks A’ and B’ in EL spectra of ML-WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction (Fig. 5-5b) demonstrates that intralayer perturbation plays a role in the formation of these HEL peaks. On the other hand, it is noted that the relative intensity of HEL peaks in BL-WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction (Fig. 5-5c) are clearly much stronger than that in ML-WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction, indicating that interlayer perturbation may also contribute to the HEL peaks (which can be further supported by temperature dependent studies, see below).

D. Investigating the origin of spin-orbital interaction in WSe\textsubscript{2}

To further probe physical mechanism governing the photon emission process in the atomically thin p-n diode, we have also conducted the temperature dependent EL studies at 25, 50 and 75 °C for both the ML- and BL-WSe\textsubscript{2}/MoS\textsubscript{2} heterojunctions (Fig. 5-5g, h), and plotted the normalized peak intensities for A and B’ peaks as a function of temperature (Fig. 5-5i). For ML-WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction, the EL intensity of all spectral peaks show a consistent decrease with increasing temperature (Fig. 5-5g,i), which is a common phenomenon in the LED devices and can be attributed to the exponential enhancement in nonradiative recombination rate with increasing
temperature. In striking contrast, temperature dependent EL in the BL-WSe₂/MoS₂ heterojunction displays highly distinct features. First, the A exciton peak in BL-WSe₂/MoS₂ heterojunction shows an unusual increase (instead of decrease) with increasing temperature. (Fig. 5-5h,i). This increase in A exciton emission may be explained by thermally decoupling neighboring layers via interlayer thermal expansion, which can induce a band gap crossover from the indirect gap to the direct one with the increasing decoupling at higher temperature. A similar thermal decoupling effect was observed in MoSe₂ by PL studies. Second, the HEL peak B’ (and A’) shows a much greater decrease with increasing temperature than that in ML-WSe₂/MoS₂, indicating the weakening electron-orbital interaction with the decoupling neighboring layers. These temperature dependent behavior are consistent seen in three devices studies and further suggests that that the interlayer perturbation plays an important role in electron-orbital interaction in WSe₂.

E. Summary

In summary, we have, for the first time, fabricated a WSe₂/MoS₂ heterojunction p-n diode with atomically thin geometry and atomically sharp interface. The scanning photocurrent measurement demonstrates that the p-n junction was formed over the entire overlapping area with a maximum photon-to-electron conversion EQE of 12%. The EL measurement allows for the identification of emission from different optical transitions. Hot electron luminescence peaks were also observed in EL spectra of WSe₂ for the first time and used to investigate the electron-orbital interaction in WSe₂. Our novel heterojunction structure offers an interesting system for investigating the microscopic nature of the carrier recombination and probing the fundamental electro-optical properties in TMD based materials, and can open up a new pathway to novel
optoelectronic devices such as spin- and valley-polarized light emitting diodes, on-chip lasers, and two-dimensional electro-optic modulators.
References


20 Xiao, D., Liu, G.-B., Feng, W., Xu, X. & Yao, W. Coupled spin and valley physics in


39 Beal, A. R., Knights, J. C. & Liang, W. Y. Transmission spectra of some transition metal


41 Wilson, J. & Yoffe, A. The transition metal dichalcogenides discussion and interpretation of the observed optical, electrical and structural properties. *Advances in Physics* 18, 193 (1969).

Figure 5-1 | Schematic illustration and band diagram of the WSe$_2$/MoS$_2$ vertical heterojunction p-n diode. a, A schematic illustration of the WSe$_2$/MoS$_2$ vertical heterojunction device show that a transferred MoS$_2$ flake on synthetic WSe$_2$ forms a vertical heterojunction. b, A schematic illustration the cross-sectional view of the WSe$_2$/MoS$_2$ vertical heterojunction device. c, The ideal band diagram of WSe$_2$/MoS$_2$ heterojunction p-n diode under zero bias.
Figure 5-2 | Structural characterization of the WSe₂/MoS₂ heterojunction p-n diode. 

- **a,** Optical microscopy image of a truncated triangular domain of monolayer WSe₂ with an inverted triangular bi-layer region at the centre. 

- **b,** The false colour SEM image of the WSe₂/MoS₂ vertical heterojunction device, with ML-WSe₂ highlighted by blue colour, BL-WSe₂ area by violet colour, MoS₂ by green colour, and metal electrodes by golden colour. The scale bar is 3 μm. 

- **c,** The PL mapping of the WSe₂/MoS₂ heterojunction device, with red colour representing the PL from MoS₂ and the green colour representing PL from WSe₂. 

- **d,** The PL spectra of synthetic ML- and BL-WSe₂ and few-layer MoS₂ flakes with the A, B exciton peaks and indirect transition I peak labelled. The intensities of BL-WSe₂ and FL-MoS₂ are multiplied by 10 times for better visibility. Inset, the B exciton peak in ML-WSe₂. 

- **e,** The cross-sectional HRTEM image of the WSe₂/MoS₂ heterojunction interface. The scale bar is 5 nm. 

- **f,** The EDS element distribution profile from the bottom to the top of Fig. 2e. The black square represent the distribution profile of W-L characteristic peaks. The green lines represent the two fitting peaks with a separation of 0.7 nm. The red line represents the fitting curve for W-L distribution profile.
Figure 5-3 | Electrical characterization of the WSe$_2$/MoS$_2$ heterojunction p-n diode. a, The $I_{ds}$-$V_{ds}$ characteristics of $n$-type MoS$_2$ FET transistor with Ni/Au (5/50 nm) contacts. b, The $I_{ds}$-$V_{ds}$ characteristics of $p$-type WSe$_2$ FET transistor with Au (50 nm) contacts. c, Gate-tunable output characteristics of the WSe$_2$/MoS$_2$ heterojunction p-n diode. d, The derivation of the p-n diode ideality factor by using a model consists of an ideal p-n diode with a series resistor. An ideality factor of 1.2 was derived with a series resistor of 80 M$\Omega$ at 0 V gate voltage (red circle), and an ideality factor of 1.3 was derived with a series resistor of 33 M$\Omega$ at -20 V gate voltage (green triangle).
Figure 5-4 | Photoresponse of the WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction p-n diode. a, Optical microscope image of the WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction. b, False colour scanning photocurrent micrograph of the WSe\textsubscript{2}/MoS\textsubscript{2} heterojunction device acquired at $V_{ds}=0$ V and $V_{BG}=0$ V under irradiation 514 nm laser (5 µW). The purple square dotted line outlines the ML-WSe\textsubscript{2} and the dark purple square dotted line outlines the BL-WSe\textsubscript{2}. The blue circle dotted line outlines the MoS\textsubscript{2} and the golden solid line outlines the gold electrodes. Photocurrent were observed in the entire overlapping junction area. c, Experimental output ($I_{ds}$-$V_{ds}$) characteristic of the vertical heterojunction device in the dark (black) and under illumination (wavelength: 514 nm; power, 5 µW). Inset, temporal response of the photocurrent generation under 514 nm illumination (10µW). d, Power-dependent EQE of the heterojunction device under 514 nm and 633 nm laser excitation wavelengths at $V_{ds}=0$ V and $V_{BG}=0$ V. A maximum EQE of 12% was observed.
Figure 5-5| Electroluminescence (EL) from the WSe$_2$/MoS$_2$ heterojunction p-n diode. a, The false colour EL image of the heterojunction device under an injection current of 100 µA. The purple dashed line outlines the ML-WSe$_2$, the blue dotted line outlines the MoS$_2$ and the golden solid line outlines the gold electrodes. b, The EL spectra of a ML-WSe$_2$/MoS$_2$ heterojunction at different injection current. c, The EL spectra of a BL-WSe$_2$/MoS$_2$ heterojunction at different injection current. d, The EL intensity as a function of injection current for both ML- and BL-WSe$_2$/MoS$_2$ heterojunction. e, The ideal band diagram of the WSe$_2$/MoS$_2$ heterojunction under small forward bias. The conduction band in MoS$_2$ is below that in WSe$_2$, the valence band in WSe$_2$ is below that in MoS$_2$. At small bias, holes can go cross the junction and inject into n-type region, while the electrons cannot go cross the junction. f, The ideal band diagram of the WSe$_2$/MoS$_2$ heterojunction under large forward bias. The conduction band in MoS$_2$ shifts upward and is higher than that in WSe$_2$, and the valence band in WSe$_2$ is below it in MoS$_2$. At large bias, both electrons and holes can go cross the junction and inject into the other side of the heterojunction. g, The EL spectra of a ML-WSe$_2$/MoS$_2$ heterojunction at different temperature ranging from 25 to 75 °C. h,
The EL spectra of a BL-WSe$_2$/MoS$_2$ heterojunction at different temperature ranging from 25 to 75 °C. The injection current is fixed at 250 µA. i, The normalized intensities of A and B’ peaks in the EL spectra of both ML- and BL-WSe$_2$ as a function of temperature.

**Figure 5-6** The AFM and Raman characterizations of synthetic WSe$_2$ domain. a, AFM image of synthetic WSe$_2$ domain. A triangular shape BL-WSe$_2$ was formed on top of a triangular shape ML-WSe$_2$. The scale bar is 5 µm. b, The height profile of the synthetic WSe$_2$ along the horizontal line shown in Fig. S1a. c, The Raman spectra of synthetic ML- and BL-WSe$_2$.

**Figure 5-7** The analysis and peak fittings for the EL spectra of both ML- and BL-WSe$_2$/MoS$_2$ heterojunction. a, The EL spectra from ML-WSe$_2$/MoS$_2$ can be fitted by 6 Gaussian peaks: A, B exciton peaks, HEL A’, B’ peaks, and the emission from indirect band gap recombination I in WSe$_2$; the A exciton peak from few-layer MoS$_2$. b, The EL spectra from BL-WSe$_2$/MoS$_2$ can be fitted by 5 Gaussian peaks: A, B exciton peaks, HEL A’, B’ peaks, and the emission from indirect band gap recombination I in WSe$_2$. c, The position of excitonic A peak in both ML- and BL-WSe$_2$/MoS$_2$ heterojunction under different injection current. A rend of red shift was observed in both ML- and BL-WSe$_2$/MoS$_2$ heterojunction due to thermal effect.