ABSTRACT

Title of Document:	HIGH FREQUENCY GENERATION BASED ON CARBON NANOTUBE FIELD-EFFECT TRANSISTORS
	Da Song, Doctor of Philosophy, 2015
Directed By:	Professor John Cumings, Department of Materials Science and Engineering

Carbon nanotubes (CNTs) are promising materials in radio frequency (RF) applications due to their high mobility, high current density and low capacitance. Over the past several years, extensive experimental and theoretical works have been focused on increasing the cut-off frequency of carbon nanotube field effect transistors (CNTFETs). However, there is limited study aiming for understanding the linearity of CNTFETs, which is an important aspect when radio frequency transistors are working in multiple frequency environments.

In this dissertation, CNTFETs are fabricated based on horizontally aligned carbon nanotubes grown on quartz substrate. DC characterization shows three conduction regions in the transfer curve of the device, p-type and n-type linear regions, and ambipolar nonlinear region. The single tone excitation measurement shows extra harmonic generations as a result of the nonlinearity of the device. Same measurement is conducted with control devices without carbon nanotubes in the

channel and confirms the nonlinearity is from the carbon nanotubes in the channel. Comparison between the 1st order harmonic amplitude and the 2nd order derivative of current with respect to gate voltage indicates that nonlinear transconductance is the cause of nonlinearity in the device. In order to understand the nonlinearity thoroughly, an elementary model based on 1D electronic transport and Drude model is built. The model can accurately predict the DC performance and nonlinearity of the device.

Taking advantage of the transitions between linear and nonlinear transfer regions, we build our CNTFETs into gate controlled radio frequency mixers. Two-tone mixing measurement shows clearly that intermodulation terms in the output spectrum are strong in ampibolar regions and suppressed to noise floor in the linear regions. We further perform passive mixing (no source/drain voltage applied) in higher frequency regime and demonstrate the generation of harmonic and intermodulation signals in the output frequency range between 75 and110GHz, which is the among the highest output frequency observed from CNTFETs to date.

.

HIGH FREQUENCY GENERATION BASED ON CARBON NANOTUBE FIELD-
EFFECT TRANSISTORS

By

Da Song

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2015

Advisory Committee:
Professor John Cumings, Chair
Dr. Vince Ballarotto
Professor Liangbing Hu
Professor Ichiro Takeuchi
Professor Romel Gomez, Dean's Representative

Acknowledgements

While writing and editing this thesis, every chapter, every figure and even the sentences I write down, they remind me that this work is not done by me, but by all the people who have helped and supported me along the way.

The first person I would like to express my gratefulness is my advisor, Prof. John Cumings. He teaches me how to think in a logical and efficient way while approaching difficult problems. His suggestions inspire me; his instructions enlighten me. He creates opportunities for me and supports from every possible aspects. The next person I need to say thank you is Dr. Vince Ballarotto, who has work with me closely. He teaches me all the test basics and discuss with me regularly. He spends a lot of time editing my writing materials and rehearsing my presentations. Without you guys, this work cannot be done.

I also received a lot of help from people working together with me. John Bavier, who is one year beyond me, teaches and helps me a lot, especially about training in the clean room. Andrew Tunnell, the previous graduate student working on the same project, prepares everything I need when he passes the project to me. He also devotes his personal time catching me up with device fabrications. Dan Hines works with me at the start of the project. He sets up the CVD furnace and trains me growth of nanotubes and SEM.

At last, I would like to thank my family. My parents encourage and support my study abroad. My wife, Guanxiong Zhang, is always there cheering me up there when I doubt myself. I would like to thank her for her enormous sacrifice to support my work.

College Park, 2014

Table of Contents

List of Tables

List of Figures

List of Illustrations

A: amplitude of ac signal

AC: alternative current

ALD: atomic layer deposition

C_g: gate capacitance

CNT: carbon nanotube

dB: decibel

dBm: decibel milliwatts

DC: direct current

E_F: Fermi energy

FET: field-effect transistor

GSG: ground-signal-ground

I_{ds}: drain current

IIP3: input third-order intercept point

IP3: third-order intercept point

LO: local oscillator

MOSFET: metal–oxide–semiconductor field-effect transistor

OIP3: output third-order intercept point

RF: radio frequency

RIE: reactive ion etcher

SEM: scanning electron microscope

SMA: SubMiniature Version A

V_{ds}: drain to source voltage

V_{gs}: gate to source voltage

Chapter 1: Introduction

Carbon serves as the backbone of life on earth. Most organic building blocks of lives contain carbon. Carbon itself is an element with 4 electrons in the outmost shell, and this allows carbon to form covalent bonds with element such as Hydrogen, Oxygen and Nitrogen, and of course, with other carbon atoms. Those covalent bonds together contribute to build up larger organic molecules.

Carbon can form different types of covalent bonds itself, i.e. sp^2 and sp^3 hybridized bonds. For sp^3 bonds, one carbon atom is connected to four other carbon atoms, forming a tetrahedron, and this is the diamond structure, where all the electrons are confined within the covalent bonds between carbon atoms. The nature pure of carbon solid is diamond. For sp^2 bonds, one carbon atom is connected to three other carbon atoms, forming a hexagon repetition two dimensional structure, like the honeycomb. In such a structure, only three out of four valent electrons are forming covalent bonds, and the other one electron is free to move, which makes such material a good conductor. Graphite is an example of such a structure. Within each layer, carbon atoms are connected by covalent bonds, and adjacent layers are held together by Van der Waals force. A single layer of sp^2 carbon is called graphene, can be pictured as a sheet of carbon atoms. And if we can roll this sheet up, what will this roll look like and what is the property of this carbon roll, or carbon tube? Such a roll is called carbon nanotube.

The first carbon nanotube (CNT) was discovered in 1991 by observation through transmission electron microscopy[1]. The cylinder normally has a length of

several microns, and in some cases, even millimeters or centimeters[2][3]. However, the radius of the cylinder is only several nanometers. Such a huge aspect ratio helps to bridge nanoscale quantum behaviors with the macroscopic devices.
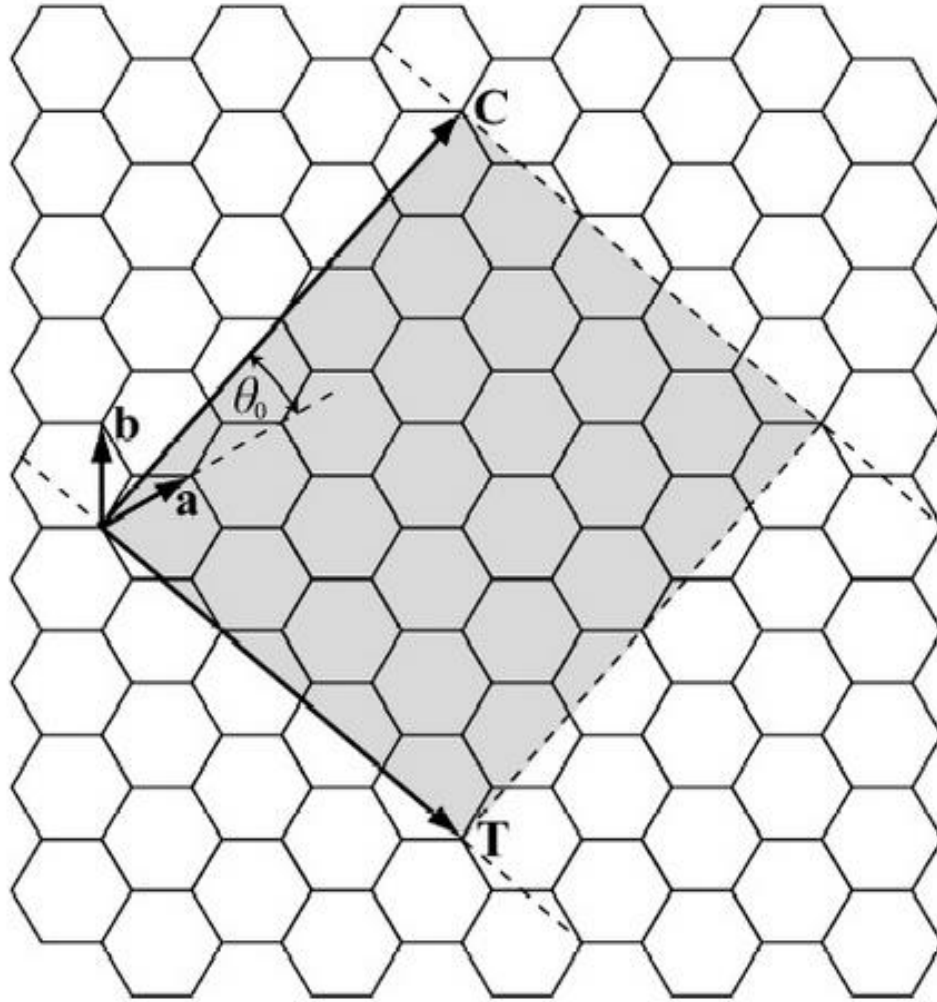
Figure 1.1 Unit cell in graphene sheet for a (4, 2) carbon nanotube. **a** and **b** are two basic vectors and **C** is the chiral vector

Carbon nanotube can be formed by only one cylinder of graphene sheet, which is single wall carbon nanotube, or several concentric cylinders, which is multiwall carbon nanotube. For single wall carbon nanotubes, in terms of how the graphene is rolled up in the circumferential direction, they can be further classified

into three categories. Define the chiral vector **C,** and **C= n a+m b** (n, m)**,** the graphene

is rolled up in the **C** vector direction. (i) when n=0, it is the zigzag nanotube (ii)when

n=m, it is the armchair nanotube (iii) all other (n, m) values corresponds to chiral

tubes. Figure 1.1 shows a (4, 2) carbon nanotube, and the shaded area is one unit cell.

According the classification above, the (4,2) nanotube is a chiral nanotube, which is

semiconducting. The chiral vector influences the electronic properties of carbon

nanotube. If (n-m) mod 3=0, the nanotube is metallic, on the other hand, if (n-m) mod

3≠0, the nanotube is semiconducting. Next section will discuss the electronic

properties of carbon nanotubes and clarify the relationship between chiral vector and

metallic/semiconducting classification.

1.2. Carbon nanotube electronic properties

In order to look at the electronic properties of carbon nanotubes, first investigate the electronic properties of graphene. The real space lattice and reciprocal lattice of graphene are shown in Figure 1.2.
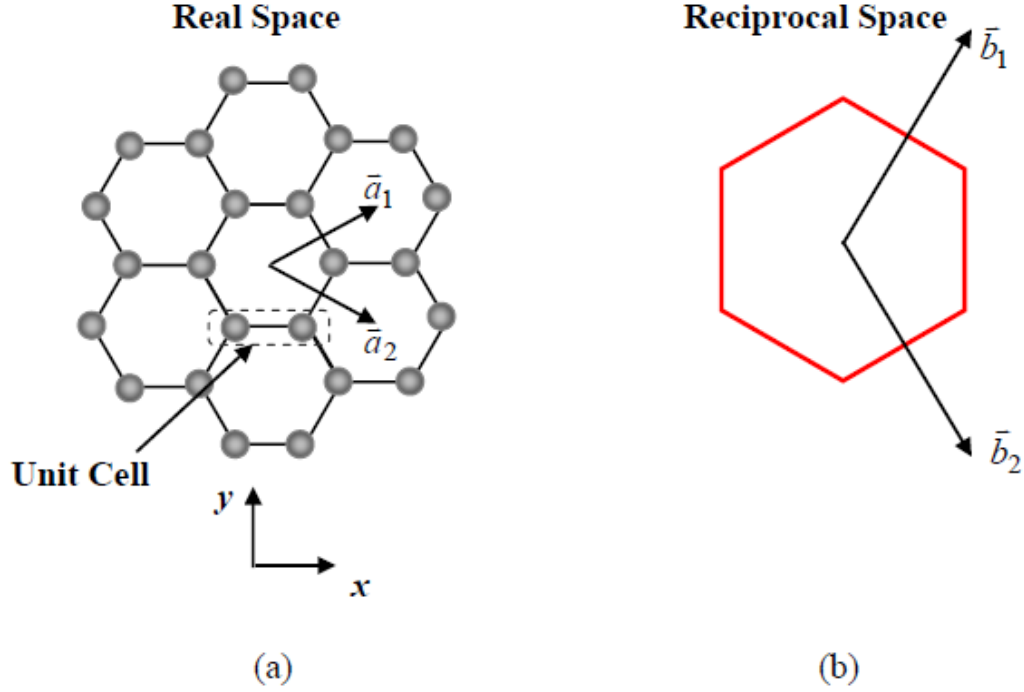
Figure 1.2 Real space lattice and reciprocal lattice of graphene

In the real space, $\vec{a}_1$ and $\vec{a}_2$ are the basic vectors and can fill up the whole plane with the help of translating vectors. $\vec{a}_1$ and $\vec{a}_2$ are expressed as

$$\vec{a}_1 = a_0 \left(\frac{\sqrt{3}}{2} \hat{x} + \frac{1}{2} \hat{y} \right)$$

$$\vec{a}_2 = a_0 \left(\frac{\sqrt{3}}{2} \hat{x} - \frac{1}{2} \hat{y} \right)$$

where $a_0 = \sqrt{3} d_c$ is the length of basis vector and $d_c \approx 1.42 \text{Å}$ is the carbon-carbon bonding length. Based on the tight bonding model[4], the dispersion relationship in graphene can be solved, and the E-k relationship is

$$E(\vec{k}) = \pm |\epsilon| \cdot \sqrt{3 + 2cos(\vec{k} \cdot \vec{a}_1) + 2cos(\vec{k} \cdot \vec{a}_2) + 2cos(\vec{k} \cdot \vec{a}_3)}$$

where $\epsilon \approx -3eV$ is carbon-carbon bonding energy, and $\vec{a}_3 = \vec{a}_1 - \vec{a}_2$. The plot of the E-k relationship is shown in Figure 1.3.
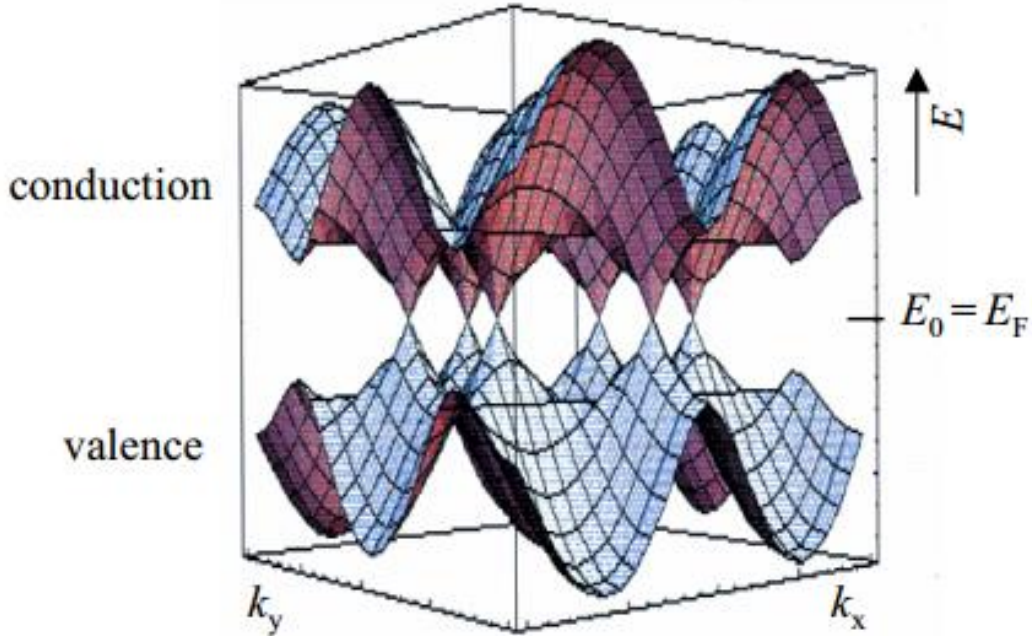
Figure 1.3 E-k relationship plotting based on tight bonding calculation

At the six corner of 1st Brillouin zone, referred to as K points, the valence band touches with conduction band but not overlapping, which makes graphene a zero band gap material. We can examine this by performing Taylor expansion for cosine function near these points and E-k relationship is

$$E(\vec{k}) = \frac{3d_c|\epsilon|}{2} |\vec{k} - \vec{k}_F|$$

5

where $\vec{k}_F$ is the Fermi vector. This illustrates that E-k follows a linear relationship near the K points, and when $\vec{k}$ equals to $\vec{k}_F$, the valence band touches the conduction band.

In terms of the band structure of carbon nanotubes, it is different from graphene's. Because of the small diameter of carbon nanotubes, the k vector along the radial direction of carbon nanotubes are quantized, which means only k vector satisfying certain conditions can be allowed. And the periodic boundary condition in this case is

$$\vec{k} \cdot \vec{C} = 2\pi j$$

where j is an integer. Further taken n-m into account, say (n-m) mod 3=q, where q is 0,1 or -1:

(1) if p=0, a line of allowed k intercepts the K point, so valence band and conduction band touch, making the nanotube metallic

(2) if p=1or -1, a line of allowed k misses the K point by $\Delta k = \pm \frac{2}{3D}$, where D is diameter of the nanotube, and the nanotube will have a band gap[5]

In Figure 1.4, the band structure, density of states and quantized graphene band structure are drawn for (a) metallic nanotubes and (b) semiconducting nanotubes. For metallic nanotubes, k vectors satisfying the boundary condition can pass the K point at the corner of 1st Brillouin zone, leading to zero band gap. For semiconducting nantoubes, k vectors misses the K point and a band gap is formed. In more detailed situation, when n-m=0, the band ganp is zero, but when n-m=3z, where z is integer,

6

there is a very small band gap in carbon nanotube band structure, but is always treated as metallic.
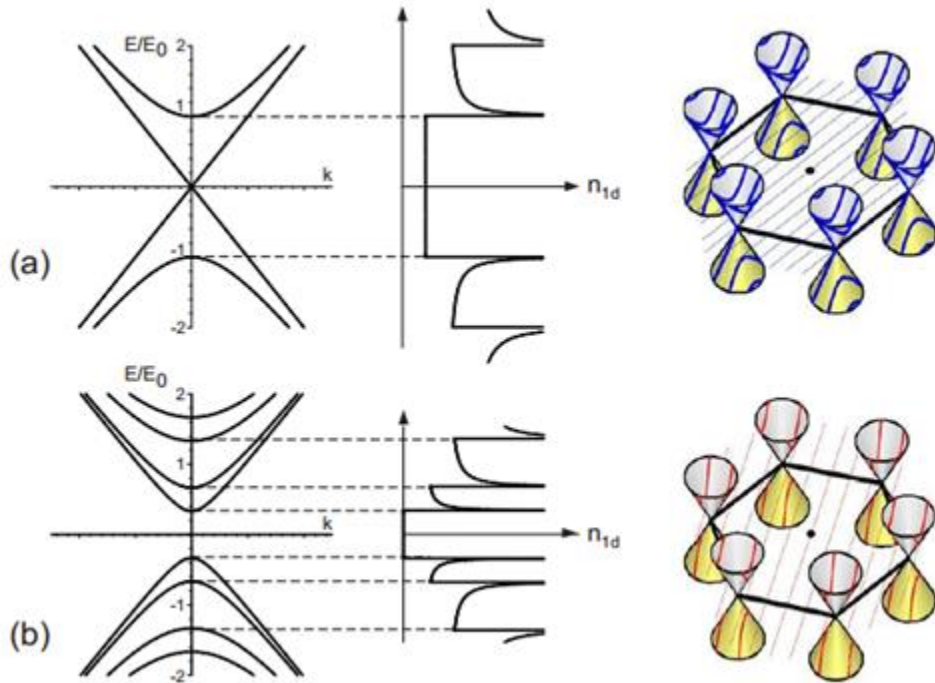
Figure 1.4 Schematic drawing of band structure, density of states and quantized graphene band structure for (a) metallic nanotubes (b) semiconducting nanotubes (image from Andrew Tunnell dissertation)

Now that we are aware of the band structure of carbon nanotubes, it is worth to mention about the electronic properties of CVD grown carbon nanotubes. CVD grwon carbon nanotubes normally have random chirality, and this end up with even distribution of (n-m). The probability of (n-m) mod 3=0 is about 1/3, and the probability of (n-m) mod 3=$\pm$1 is 2/3. So one third of the CVD grown nanotubes are metallic and two thirds of the nanotubes are semiconducting. It is challenging to grow either metallic or semiconducting nanotubes solely, and removal of metallic nanotubes is also an unsettled topic. However, metallic nanotubes will not show gate control behavior, which means they are not going to contribute to nonlinearity of the

device in a carbon nanotube field-effect transistor (CNTFETs). So we choose not to remove the metallic nanotubes and deal with the mixsure of both metallic and semiconducting nanotubes. By doing so, we are still studying the nonlinearity behavior caused by semiconducting nanotube..

1.3. Carbon nanotube high frequency applications

Ever since the discovery of carbon nanotubes, these small diameter hollow tubes have drawn greater and greater attention due to their interesting properties in a number of regions. Carbon nanotube is a very stiff material with high Young's modulus[6], and it is also an efficient thermal conductor, whose thermal conductivity is near 3000W/mK[7], which is more than 7 times than that of copper. However, in this dissertation, we will focus more on another property that enables extensive study of carbon nanotubes, the extraordinary electronic properties.

Among the carbon nanotube family, single wall carbon nanotubes show some superior properties. One outstanding aspect is they can carry very high current density, as high as $\sim 4 \times 10^9 \text{Acm}^{-2}$, which is three orders of magnitude higher than a typical metal such as copper or aluminum[8][9]. On the other hand, they also have high carrier mobility of $\sim 10,000 \text{cm}^2 \text{V}^{-1} \text{s}^{-1}$, which is better than that of silicon[10][11]. Such high carrier mobility encourages people to use carbon nanotubes as the active material for high frequency devices. Since the carrier mobility is closely related to the drift velocity, and the higher the drift velocity is, the quicker the device can react to the AC signals applied on it. Thus, higher carrier mobility

should enable devices to work at higher frequency, especially when the channel length of the transistor is scaled down[12].

Although carbon nanotubes can carry a high current density, based on its small diameter, using single carbon nanotube to build electronic device has limitations in terms of practical use and measurement due to its large output impedance which does not match most measurement apparatus and the small output current. Most current carbon nanotube based devices consist of tens or hundreds of nanotubes in the device channel, enabling practical output currents[13].

For the analysis of devices working in high frequency region, there are basically two aspects. The first aspect relates to the amplitude of output signal, where characterizations mainly focus on what type of gain is possible and at what frequency region these gains exist. Two different definitions of gains are widely used to characterize the frequency response of carbon nanotubes, the current gain H_{21} and the Mason's unilateral gain U. The current gain H_{21} is defined as the ratio of output current to input current, while the unilateral gain U is the power gain under conjugate impedance matching. For a field-effect transistor, two figure of merits are more often used instead of the amplitude of the gain, they are the cut-off or transition frequency f_T, at which the current gain falls to unity and the maximum frequency of oscillation, at which the unilateral gain falls to unity. Based on the assumption of ignoring the parasitic capacitance between source/drain and gate, Rutherglen et al. calculate the intrinsic cut-off frequency of a CNTFET[14],

$$f_{T,intrinsic} = \begin{cases} \dfrac{\mu(V_{gs} - V_T)}{2\pi L^2} & \text{L is large} \\[2ex] \dfrac{v_{sat}}{2\pi L} & \text{L is small} \end{cases}$$

9

where μ is the mobility, V_{gs} is the voltage between gate and source, V_T is threshold

voltage and is defined as $V_T = V_{gs} - V_{ds}$ in the current saturation regime, V_{ds} is the

voltage between source and drain, v_{sat} is the saturated drift velocity at large electric

field (small channel length), and L is the channel length of the device. We notice that

in either of the two limits, decreasing the channel length will lead to increase of cut-

off frequency, and this is the motivation of several groups to aggressively down-size

the channel length of CNTFET to achieve higher cut-off frequency[15][16], where

techniques such as e-beam lithography[17][18] and self-aligned source/drain[19][20]

are used. Theoretical projections are also focused on the limitations of CNTFETs

radio frequency performance[21][22][23], which covers short channel length that is in

the ballistic regime[24]. Our devices usually have a channel length from 2μm to 5μm,

so the cut-off frequency is relatively low, but this will not be an issue here as we plan

to investigate the linearity of the devices working at radio frequency.

The linearity of the device comes to play an important role, and is becoming

increasingly significant a figure of merit when we want our device to work in a

multiple signal environment. There have been efforts to directly observe the nonlinear

behavior of carbon nanotubes, and some are listed in Table 1.1. The maximum input

frequency and the type of output signal are also listed in the table.

Table 1.1.   Direct observation of nonlinear behavior of carbon nanotubes

Year	Author	Maximum Input Frequency (GHz)	Output signal	Type of detection
2005	Rosenblatt[25]	50	DC	DC mixing current
2006	Pesetski[26]	23	10 kHz	DC mixing current
2008	Cobas[27]	18	DC	Diode rectification
2011	Wang[15]	1	3 GHz	Higher order harmonics
2014	Wang[18]	10	20GHz	Higher order harmonics

Based on the table, we notice that there is a need for a characterization of the nonlinear behavior of carbon nanotube over a large frequency range using consistent detection methods. However, based on the range of commercially available signal analyzer, detection of higher order harmonics beyond the range of 50 GHz is not economical. So here we propose using a measurement setup consisting of a signal analyzer and an external mixer to characterize the nonlinear behavior of carbon nanotubes. By measuring the higher order harmonics resulting from single tone harmonic distortion or two-tone mixing, characterization of the nonlinear behavior of carbon nanotubes FETs, will be done from 1 GHz up to 110 GHz. The possible reasons for the nonlinear behavior will be discussed based on both experimental data and theoretical calculation. Moreover, we will utilize the nonlinearity of CNTFETs to build novel radio frequency circuits, such as frequency doubler and mixer.

1.4. Overview of the dissertation

The goal of our research is to fabricate carbon nanotube field effect transistors (CNTFETs) based on horizontally aligned carbon nanotubes, study the nonlinearity of CNTFETs by measurement and interpretation of harmonic generation, and take advantage of the nonlinearity to build radio frequency circuit components (frequency doubler and mixer). In Chapter 2, we introduce the CVD growth of horizontally aligned carbon nanotubes on quartz substrate, go over the fabrication process of CNTFETs, and study the DC characterization of our devices. In Chapter 3, we describe the experimental setup of single tone excitation measurement and the observation of extra harmonic generation in the output spectrum. Harmonic generation theory based on small signal model is illustrated and the comparison between 1st order harmonics and 2nd derivative of current with respect to gate voltage confirms the connection between DC and AC performance. A frequency doubler based on CNTFETs working at low frequency is demonstrated. In Chapter 4, we build up our elementary model of electronic transport in CNTFETs based on a combination of Drude model and 1D electronic transport. We further compare the nonlinearity based on the model with the amplitude of 1st order harmonics from single tone excitation measurement. In Chapter 5, we introduce the two-tone mixing measurement and generic mixing theory. Based on the unique control of gate voltage over the nonlinearity of the device, we are able to achieve controllable mixing using CNTFETs. The first observation of CNTFETs mixing at 75-110GHz range is presented. Chapter 8 summarizes and concludes the dissertation, and suggests some possible further work.

Chapter 2: Device fabrication and DC operation

2.1. Growth of horizontally aligned carbon nanotubes

1. Advantage of using horizontally aligned carbon nanotubes as channel materials for CNTFETs

It has been clearly pointed out that, in order to increase the radio frequency figure of merits, one important aspect is to increase the gate capacitance contribution of the total capacitance and decrease parasitic capacitance related to gate/drain and gate/source. In order to achieve such condition, the most effective way is to increase the density of carbon nanotubes in the channel. Because the gate capacitance is proportional to the number of nanotubes in the channel, however, the parasitic capacitance is mostly from the device structure. So by increasing the density of carbon nanotubes in the channel, the gate capacitance per nanotube is kept still, but the parasitic capacitance per nanotube decreases. To evaluate the frequency response of a transistor, current gain H_{21} is usually an important figure of merit. It is defined as the output power divided by the input power. The cut-off frequency f_T is the frequency at which H_{21} falls to unity, the f_T for a CNTFET has been derived as

$$f_T = \frac{g_m}{2\pi} \frac{1}{\left(C_{gs}+C_{p,gs}+C_{p,gd}\right)\left(\left(R_{p,s} + R_{p,d}\right)g_d + 1\right) + C_{p,gd}g_m(R_{p,s} + R_{p,d})}$$

The parameters in the equations are labeled out in Figure 2.1, where g_m is the transconductance, C_{gs} the intrinsic gate capacitance, and g_d the channel conductance. $C_{p,gs}$ and $C_{p,gd}$ are the gate-source and gate-rain parasitic capacitances, $R_{p,s}$ and $R_{p,d}$

are parasitic resistances for the source and drain, and R_{gate} is the resistance of the gate electrode. The green dash line highlights the intrinsic part of the device.
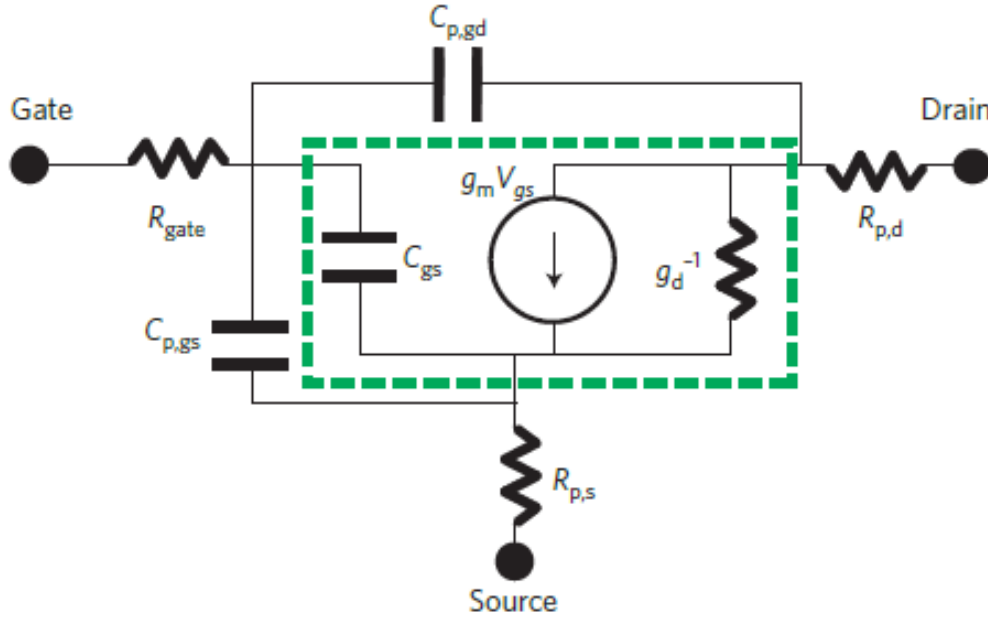
Figure 2.1 A Small circuit model based on a CNTFET. g_m is the transconductance, C_{gs} the intrinsic gate capacitance, and g_d the channel conductance. $C_{p,gs}$ and $C_{p,gd}$ are the gate-source and gate-rain parasitic capacitances, $R_{p,s}$ and $R_{p,d}$ are parasitic resistances for the source and drain, and Rgate is the resistance of the gate electrode. The green dash line highlights the intrinsic part of the device.

The important information we get from the equation is that, increase in $C_{p,gs}$ and $C_{p,gd}$, which are the gate-source and gate-rain parasitic capacitances, will lower the cut-off frequency. Increase the density of carbon nanotubes not only decreases the parasitic capacitances per nanotube, it also increases g_m, and both actions will increase f_T.

2. CVD growth of horizontally aligned carbon nanotubes on quartz substrates

The growth of horizontally carbon nanotubes was first demonstrated by Rogers' group[28]. Such alignment can be achieved on different substrate, such as

quartz and sapphire[29], with different metals as catalyst[30][31]. Also, the carbon feeding gases have alternative choices, methane, ethanol and methanol can all serve as carbon source[32]. Liu group from Duke[33], using copper as catalyst and mixture of ethanol and methanol as carbon source, has demonstrated preferable growth of semiconducting carbon nanotube arrays. In terms of increasing the density of carbon nanotube arrays, physical methods including sequential CVD and transferring after growth have been explored, and the highest density achieved can reach the value of tens of tubes per micron. On the other hand, chemical methods such as introducing sulfur into the system for Fe catalyst assisted reaction have also been shown to successfully improve the density from ~2/µm to ~8/µm. Growth in place using CVD enables patterning the catalyst in the desired region, and this enables the scalable growth of carbon nanotubes with good alignment in designated region, which is compatible with the modern semiconductor industry process and also enables later transferring to alternative substrate like silicon[34]. The disadvantages of this technique includes high temperature process up to ~850℃ and the coexistence of both metallic and semiconducting carbon nanotubes, and the second problem has attracted a lot of attention since it is difficult to remove metallic carbon nanotubes without influencing the density and integrity of semiconducting nanotubes.

As is mentioned, one of the problem using carbon nanotubes to substitute semiconducting materials in electronic device is that CVD grown nanotubes always have both metallic and semiconducting types and the ratio is about 1:2[35][36]. Attention has been paid to separate metallic and semi-conducting carbon nanotubes. Solution based separation can lead to large scale fabrication of carbon nanotube

network based electronics, and these methods include centrifugation and DNA assisted separation[37]. For carbon nanotubes grown locally on quartz substrate, the separation mainly incorporates plasma etching[38] or electric burning[39][40], and some groups also proposed in-situ etching with –OH groups[41] or even preferential growth of semiconducting nanotubes[42]. Early research demonstrates selective etching of metallic type carbon nanotubes using methane plasma. However, such etching method is diameter dependent and also will etch semiconducting nanotubes. Recently, Rogers' group use thermocapillary flows to create openings in thermally sensitive resist above metallic nanotubes while all semiconducting nanotubes are gated off, and after the metallic tubes are exposed, oxygen plasma is used to remove the metallic tubes[43]. This method reach the goal of selective removing metallic tubes, however, the experimental process is very complicated and hard to reproduce. Previously, A. Tunnell has proposed a way of quickly removing metallic tubes from arrays of CVD grown carbon nanotubes[44]. The idea is to use a silicon wafer with thermal grown oxide as temporary top gate, even though there is air gap between the dielectric layer and the carbon nanotubes, it can still gate-off semiconducting nanotubes to some extent. Although the on state current of the transistor is lowered, the on-off ratio is increased by an order of 10^4. However, in our work, the main goal is to investigate the nonlinear behavior of carbon nanotubes, where the metallic nanotubes actually do not contribute much, so removal of the metallic nanotubes is not concentration.

The growth of horizontally aligned carbon nanotubes can be described as following. We start with stable temperature cut single wafers (42.75°) from

Hoffman Materials. After cutting into 12mmx15mm pieces, the quartz substrates are put into a 2 inches process tube and annealed in air at 880℃ for 12 hours. The annealing step is a prerequisite for the alignment of carbon nanotubes due to the influence of surface morphology of the quartz substrate.

After annealing, Fe thin films are deposit on the quartz substrate in regions defined by photolithography with a thickness of about 0.5 nm. This will make the nanotubes having a diameter between 1nm and 2 nm[45][46]. The sample is then annealed in the open air at 690℃ for 10 min, allowing formation of catalyst nanoparticles. After cooling to room temperature, a hydrogen flow of 300sccm is introduced into the processing tube for 45 min while the temperature is increased from room temperature to 790℃. After stabilizing at 790 ℃for 5min, the reaction gases, 1000sccm CH4, 120sccm H2, and 5sccm Ar bubbled through H2O are introduced into the system for 1 hour. The purpose of water is to increase the growth speed of carbon nanotubes[47][48][49]. This process is also shown in Figure 2.3. (a) (b) and (c).

The SEM images of as grown nanotubes are show in Figure 2.2. Figure 2.2(a) shows the whole catalyst area for an individual CNTFET. After putting down source and drain, the two regions of carbon nanotubes in the middle of the figure will become the double channel of the CNTFET. In Figure 2.2 (b) is the zoomed in SEM image, carbon nanotubes are very well aligned especially in the middle area between catalyst regions, while near the catalyst region, the alignment is not very good. Based on this, our channel will be put down right in the middle between the two catalyst line

17

regions. The density of carbon nanotubes in the array is about 2~3 nanotubes per micron. Because we are not pursuing high cut-off frequency, the density of 2~3 tubes per micron is good enough. Actually, if the carbon nanotube density is too high in the channel, say over 50 or even 100 tubes per micron, we can no longer treat the electronic properties as the sum of individual nanotubes. As the density increases, the interaction between carbon nanotubes becomes a more and more important factor. The most important aspect is that the capacitance between gate and nanotube arrays will not have simple relationship with individual nanotube gate capacitance, which will make the modeling of electronic transport in CNTFETs more difficult.
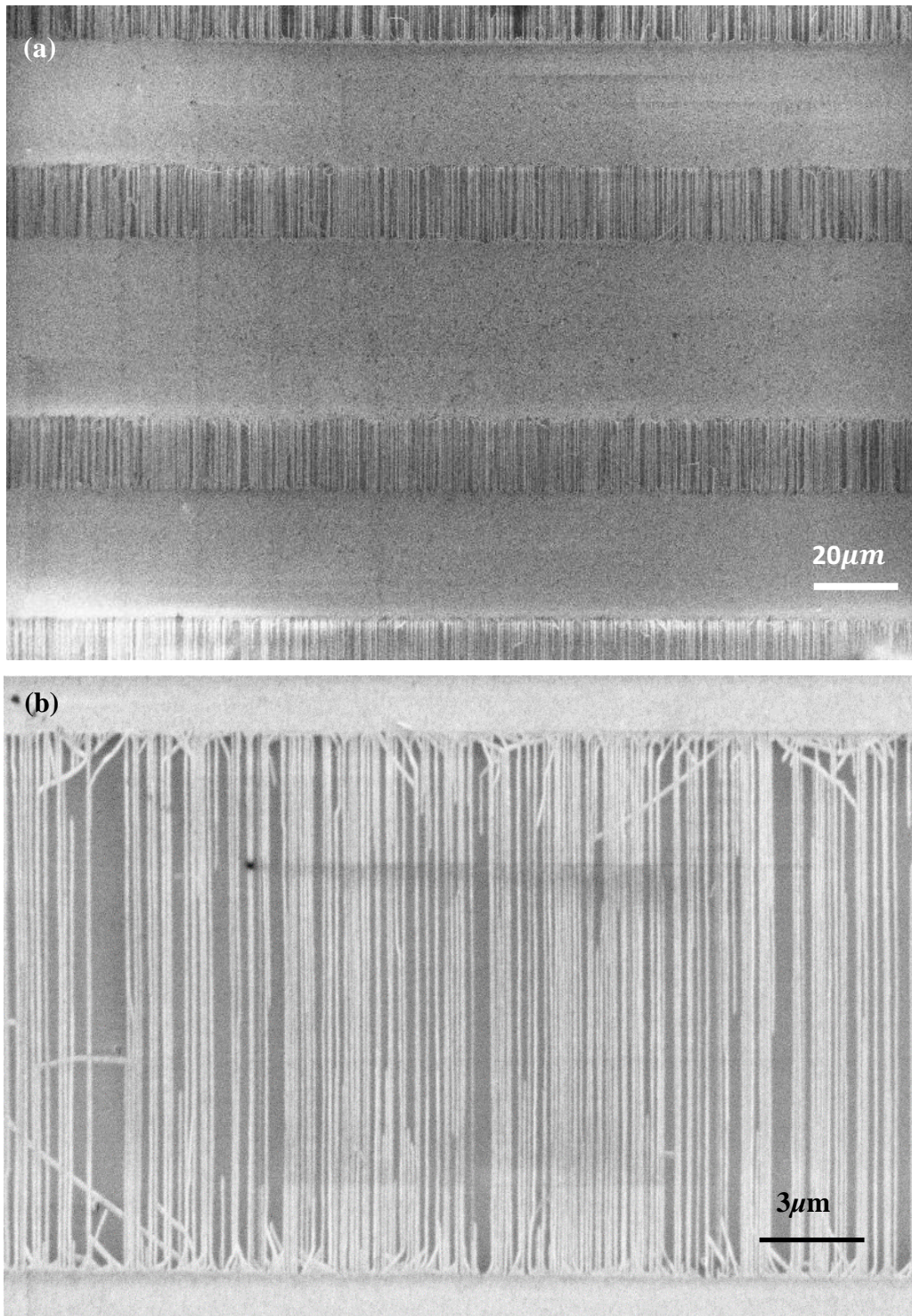
Figure 2.2 SEM image of as grown carbon nanotubes on quartz substrate (a) shows a large area view and (b) shows smaller area, with a carbon nanotube density of 2~3 tubes/micron
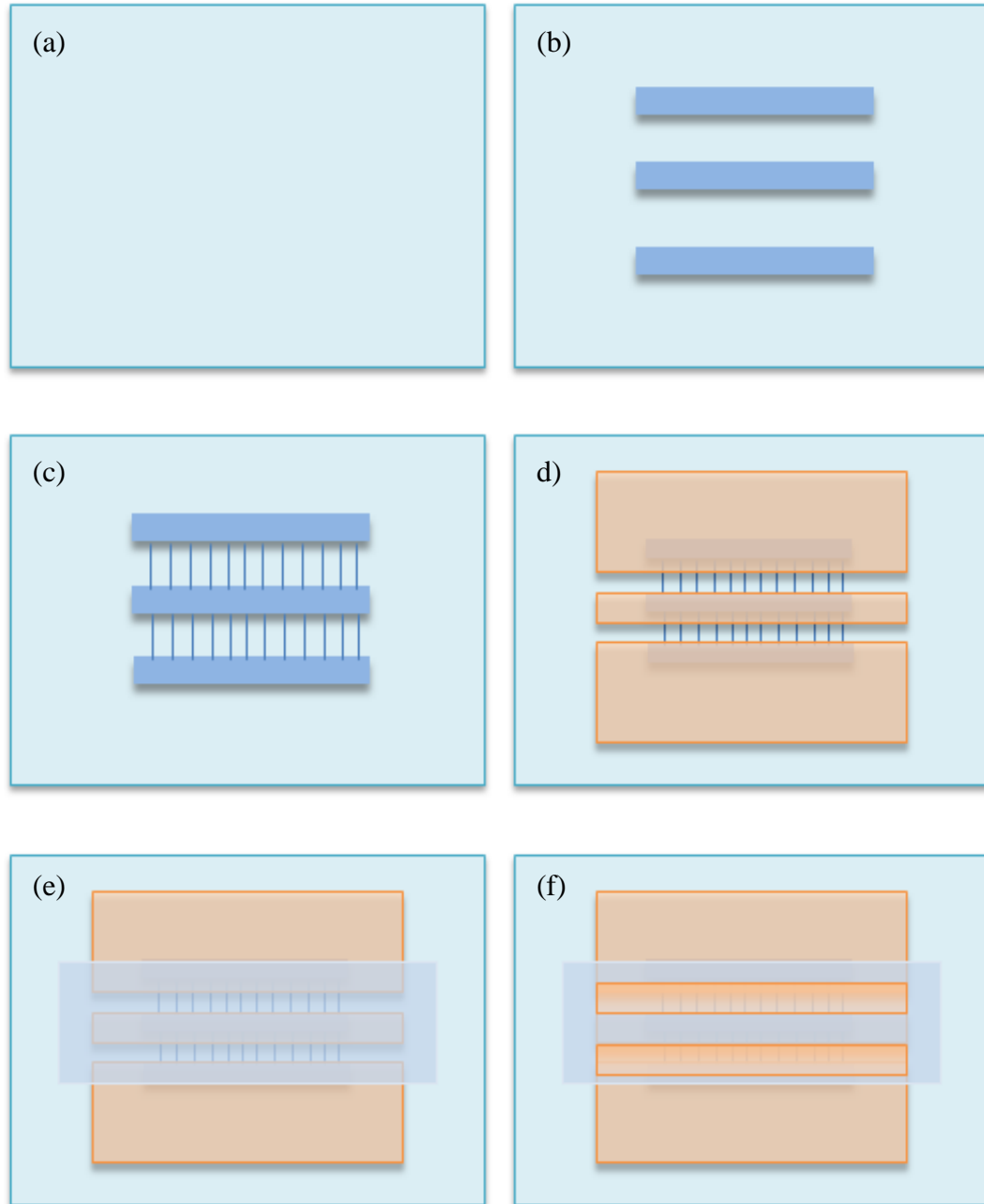
Figure 2.3 Schematic drawing showing the device process procedure. (a) start with ST cut quartz wafer and annealed at 880℃ for 12hrs (b) patterning of Fe thin film (0.5nm) (c) anneal catalyst to form Fe nanoparticles and grow carbon nanotubes at 790℃ with 1000sccm CH_4, 120sccm H_2 and 5sccm Ar flow through water bubbler as reaction gas (d) photolithography of source and drain with Ti/Pd(e) ALD deposition of Al_2O_3 with following HF etch to define dielectric layer (f) photolithography of gate using Ti/Pd

2.2. Device fabrication

The device structure chosen to study the radio frequency behavior of carbon nanotubes is usually based on a metal–oxide–semiconductor field-effect transistor (MOSFET) structure. Since the active channel material consists of carbon nanotubes, they are generally called carbon nanotube field-effect transistors. Both back gate and top gate structures have been used[50], but top gate structure inherently increases the robustness of the device and prevents the issue of environment disturbing such as gas absorption in the air[51]. Another advantage of the top gate structure is the increase in transconductance, which indicates an overall better control of gate over the nanotubes, and this better gate control can be attributed to larger gate capacitance[52]. In our device, we use Atomic Layer Deposition (ALD) to deposit Al_2O_3 as the dielectric layer of our device, the thickness of Al_2O_3 is about 70nm, and the electric constant is around 8. These two facts can contribute to higher transconductance as well as better gate control.

Figure 2.4 shows schematic drawing of the structure of a top gate CNTFET. The reason that double channels are used is because this will enable the use of a ground-source-ground probe geometry which is readily available for high frequency measurements. For a majority of our measurement, the input RF signal is through the gate, while the output signal is read from the drain.
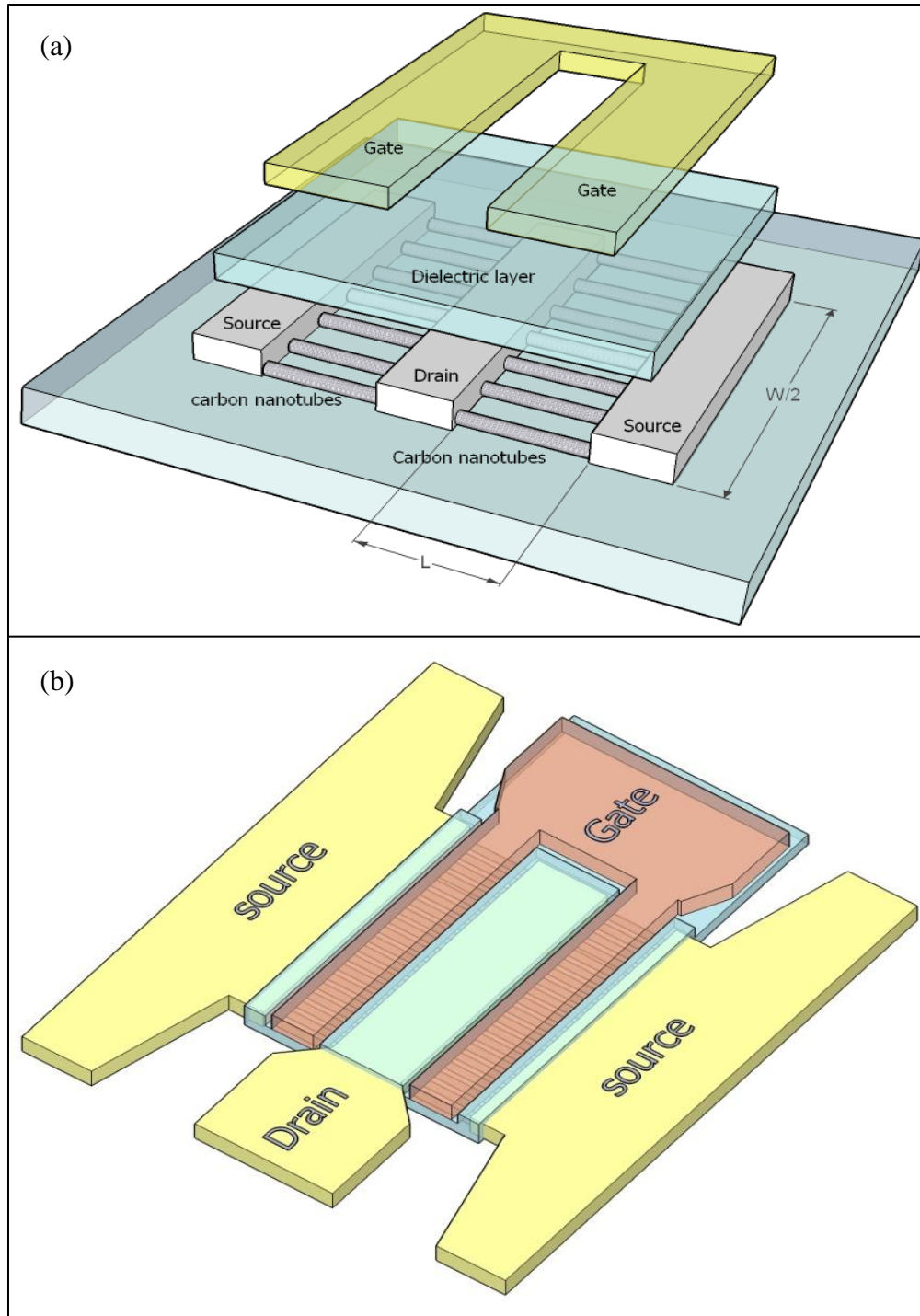
Figure 2.4 Schematic drawing of a carbon nanotube field effect transistor. (a) the layer by layer device process procedure from bottom to top (b) the completed device

Figure 2.3 shows the fabrication process step by step. We begin with the ST-cut quartz substrate. After the substrate being annealed in the open air at 880ºC for 12hrs, a 0.5nm thick Fe thin film is deposited on the substrate, the Fe thin film deposited area is defined by photolithography. Within the same step, we also put down the alignment marker with 30nm Ti for later alignment requirement. The sample is then heated at 690 ºC in the open air for 10min, leading to the formation of Fe nanoparticle in the defined area. After cooling down to room temperature, CNT growths are conducted at 790 ºC with CH_4, H_2 and H_2O as reaction gas. Next step is deposition of source and drain. For a typical Pd contact device, 1nm Ti and 50nm Pd is deposited onto photolithography defined area using e-beam evaporation. After liftoff, the next step is using Reactive-ion Etching (RIE) to remove the carbon nanotubes outside the source/drain channel. The Al2O3 dielectric layer is deposited using ALD at 160 ºC, followed by 100:1 HF dip wet etching to expose the source and drain. The last step is defining gate using photolithography, with 5nm Ti and 50nm Pd.

Figure 2.5 shows the optical images of CNTFETs and SEM image of the carbon nanotubes in the channel. Figure 2.5 shows the optical image of an individual carbon nanotube field effect transistor with channel length of 4μm and gate length 4μm. (b) shows arrays of CNTFETs and we have over 60 individual devices with different geometry parameters on a single piece of quartz substrate. (c) shows the SEM image of carbon nanotubes in the 4μm channel of the device. Carbon nanotubes in the channel show good alignment and keep the high density from as grown nanotubes.
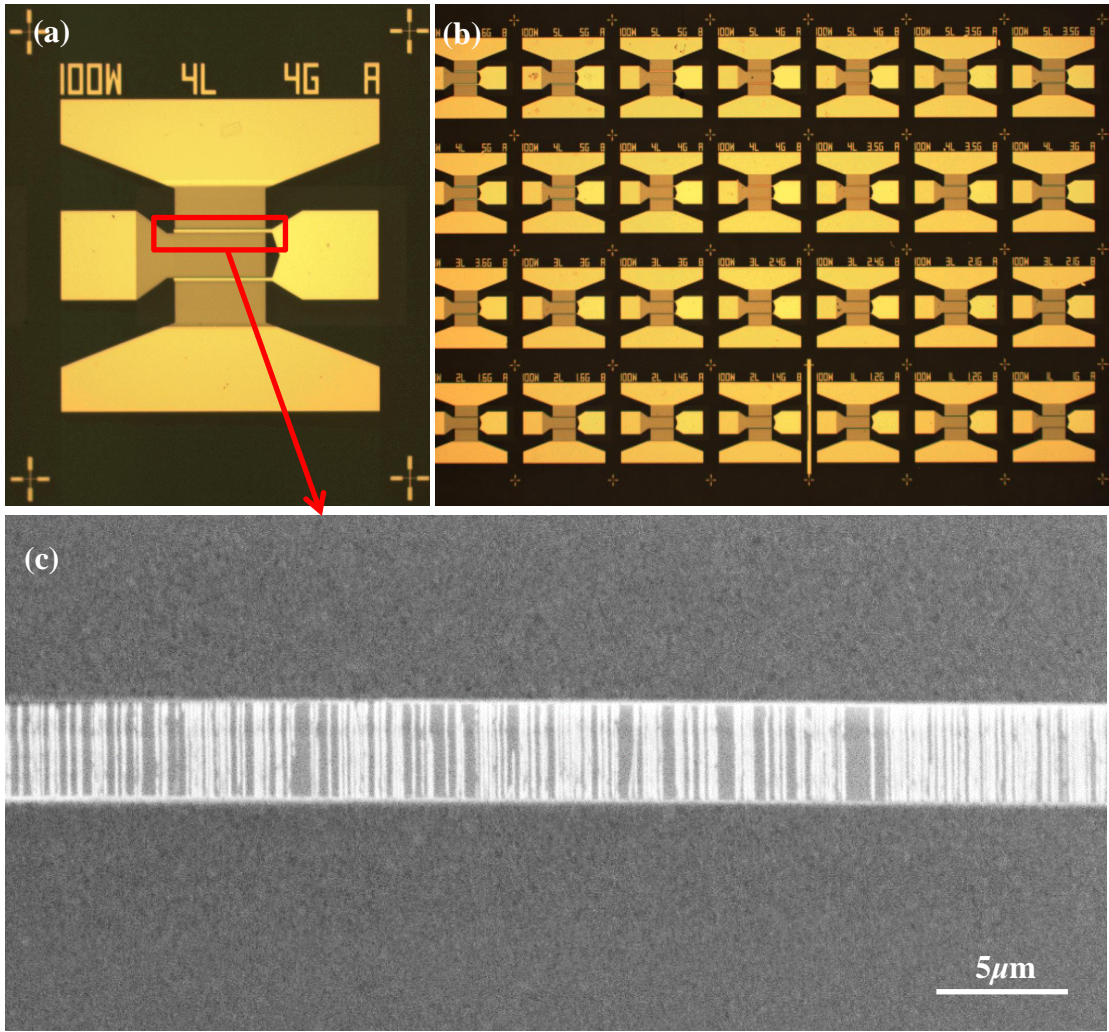
Figure 2.5 (a) Optical image of an individual carbon nanotube field effect transistor (b) Optical image of carbon nanotube field effect transistor array (c) SEM image of the carbon nanotubes in a 4μm channel device

2.3. DC operation of field effect transistors
1. Field effect transistor operation

A field effect transistor operation (FET) controls the flow of carriers from the source to drain by affecting the "conductive channel" created and influenced by the bias between gate and source. There are several kinds of FETs based on how the gate is physically built or connected to the conductive channel. In our work, the devices have MOSFET structure, this structure is widely used to incorporate low dimensional materials as the channel materials in FETs, for example, MOSFET structure graphene transistors[53] and MoS_2 transistors[54].

In our CNTFETs, the semiconducting materials are semiconducting nanotubes in the channel. The gate voltage will move the Fermi level of semiconducting nanotubes. Based on the diameter and chirality of the nanotubes, and the metal contact as well as the dielectric materials (Al_2O_3), carbon nanotubes show mostly p-type conduction, with less significant conductivity on the n-side. However, larger diameter nanotubes can lead the device to have relatively symmetric ambipolar characteristics. Most of our devices show p-type conduction at negative gate voltage, as the gate voltage increase, ambipolar conduction shows up before the devices turn into n-type conduction with a smaller conductance than p-side. This means the Fermi level of our device lies closer to the valence band of the semiconducting tubes, which is usually seen with devices exposed to air and with Al_2O_3 as dielectric materials. However, the fact that metallic carbon nanotubes also bridge the source and drain cannot be neglected. The metallic tubes have no gate effects, but will still contribute to conductivity between source and drain whenever there is a voltage difference.

2. MOSFET DC operation classic equations

25

The classic MOSFET operation theory is based on bulk semiconductor transistors. Although the transport mechanics of bulk semiconductor is different from 1D semiconductor, which is carbon nanotube in our discussion, we will use the terminology and basic DC operation equations as a start of the description of CNTFETs DC operation. A MOSFET (n-type) can be operated in three different regions:

The first operation region is the subthreshold, or cutoff region, this happens when V_{gs} is smaller than V_{th}, where V_{gs} is the gate to source bias and V_{th} is the threshold voltage. The current in the channel is very small and sometimes are treated as the off-state of the transistor. However, taking a more accurate prospective, there is weak inversion current due to the high energy carriers, and this will result in a subthreshold current exponentially related to gate-source voltage.

The second operation region is the triode mode, or linear region, where $V_{gs} > V_{th}$, and $V_{ds} < V_{gs} - V_{th}$. Here V_{ds} is the source to drain bias. The transistor in active mode at this region, and the current in the channel is a function of both V_{ds} and V_{gs}. A classic equation to describe the current-voltage relationship is

$$I_{ds} = \mu_n C_{ox} \frac{W}{L} \left[(V_{gs} - V_{th})V_{ds} - \frac{V_{ds}^2}{2} \right]$$

Where μ_n is the carrier mobility, C_{ox} is the gate capacitance per area, W is the channel width and L is the channel length. We also need to introduce two parameters that will be mentioned a lot, the transconductance g_m and dynamic conductance G_d,

$$g_m = \frac{\partial I_{ds}}{\partial V_{gs}} \qquad G_d = \frac{\partial I_{ds}}{\partial V_{ds}}$$

In the linear region, I_{ds} is changing linearly with gate voltage and g_m is constant.

The third operation region is the saturation region, where $V_{gs} > V_{th}$, and

$V_{ds} \geq V_{gs} - V_{th}$. In this region, the approximate equation is as following

$$I_{ds} = \frac{\mu_n C_{ox}}{2} \frac{W}{L} (V_{gs} - V_{th})^2$$

where I_{ds} is following the square law.

In terms of our CNTFETs, it can be operated in all three regions. However, with the existence of metallic nanotubes, it is hard to determine the threshold voltage and the turning point between linear region and saturation region. Further, these operation equations are based on bulk or planar silicon MOSFETs, the electronic transport mechanism is different from the 1D electronic transport in carbon nanotubes. It has been shown that based on the physical parameters extracted from a single wall carbon nanotube FET, the classical MOSFET operation equation is applicable only in small voltage region, and the performance of the device then violates the square law at larger voltage[55]. This is the reason why we are going to develop our own model which includes 1D transport characteristics.

3. Nonlinear behavior in CNTFETs

The nonlinear behavior in CNTFETs has drawn a lot of attention since it is crucial for the potential application of high frequency amplifier and mixers. For high frequency amplifier, linearity is valuable so that signals with information can be transmitted with higher power and do not interrupt with each other or with the higher order harmonics generated. On the other hand, nonlinearity is the prerequisite for a mixer, because linear IV curve will not offer any higher order terms in the Taylor expansion.

So for different high frequency circuit components, the desired linear (nonlinear) behavior of the device also differs.

The most noticeable nonlinearity source of CNTFETs is the nonlinear transconductance, in other words, it is the nonlinear response of conductivity of carbon nanotubes to the change of gate to source voltage that introduces the nonlinear behavior. It has been proposed that, for very shot channel length CNTFET, shorter than the mean free path of carrier, there is inherent linearity in the current/voltage relationship[56]. However, the requirement to achieve the ballistic transport is hard to achieve. Although carbon nanotubes have long mean free path, which is larger than 100nm[57], it requires e-beam lithography commonly to achieve such a short channel length, and this means no scalable fabrication of devices. And even though we can set our step into the ballistic region, there are still other factors that will hinder us from taking advantage of the high linearity, for example, nonlinear quantum capacitance of carbon nanotubes[58]. Based on what we mentioned about building high frequency transistor, linearity is one issue, and the other issue is getting gain out of the transistor at higher frequency. But the influence of parasitic capacitance increases when we get shorter channel, since the gate capacitance from the coupling between nanotubes and gate is smaller, however, the parasitic capacitance is not influence by the channel length, and is more about the width of the channel. So if we want to pull out gain from CNTFETs, it requires novel device structure that mitigates the effect of parasitic capacitance.

In our device, the channel length is from $2\mu m$ to $10\mu m$, and that means we are still in the acoustic phonon scattering region. So there are probably nonlinear

behaviors arousing from the scattering of carriers, which will be taken into account when we try to model the carrier transport in CNTFETs. The dielectric layer thickness of the devices is around 70nm, and the capacitance from the geometric gate coupling is much smaller than the quantum capacitance. So the dominating capacitance is still the geometric gate capacitance.

When it comes to nonlinearity contributions from the source/drain, one thing needs to be taken into consideration is the Schottky barrier between the semiconducting nanotubes and the metal contacts[59][60][61]. There is also research work studying the transition between Schottky and Ohmic contact to CNTs[62]. Devices have been built based on individual carbon nanotube with schottky contacts[63]. For our devices, we have multiple source/drain metal selections. Ti/Au and Pd are the most common contacts we use. It has been proven that Pd can actually form ohmic contacts to carbon nanotubes due to its high work function[64]. Even Ti/Au can form small Schottky barrier and sometimes ohmic contacts to carbon nanotubes. We also made devices with different metals contacts as source and drain, exploring Schottky barriers effect on the nonlinearity of the device.

4. DC characterization of CNTFET

Figure 2.6 (a) shows the transfer curve of a CNTFET. The device is a $4\mu m$ channel length and $100\mu m$ channel width CNTFET with gate length of $3.5\ \mu m$. The voltage between source and drain is 1V. In the transfer curve, we can identify three different conduction regions. From negative gate voltage to positive gate voltage, we have p-type conduction region, ambipolar conduction region and n-type conduction

29

region respectively. The transconductance, which is the 1st derivative of current with respect to gate voltage, is ~0.3mS in p-type conduction region and ~0.1mS in the n-type conduction region. The transistor cannot be fully turned off, the on/off ratio is larger than 2, and can further be improved if we bias gate voltage more negative. The main reason that the transistor cannot be turned off is because of the existence of metallic nanotubes in the channel. And the portion of metallic nanotubes is about one third, which should give an on/off ratio around 3.

Figure 2.6 (b) is a schematic drawing of the transition from p-channel to ambipolar channel then to n-channel. The reason we can tune the transistor from p-type conduction channel into n-type conduction channel is mainly because of the small band gap of the nanotubes in the CNTFET. Although the band gap of carbon nanotubes depends on the chirality nanotubes, it generally reversely related to the diameter of carbon nanotubes. We can describe the transition between different conduction regions as following: (1) the carbon nanotube in the channel is p-type doped due to oxygen absorption, so the Fermi level is close to the valence band. When the gate voltage is negative, Fermi level move downwards below valence band, this encourages p-type conduction. (2) When the gate voltage moves to less negative region, the Fermi level moves into the middle of the band gap, due to the small band gap of carbon nanotubes, both holes and electrons can be responsible for carrier transport, and this is the ambipolar region. (2) At more positive gate voltage, Fermi level leaves the band gap and moves into conduction band, electron becomes the
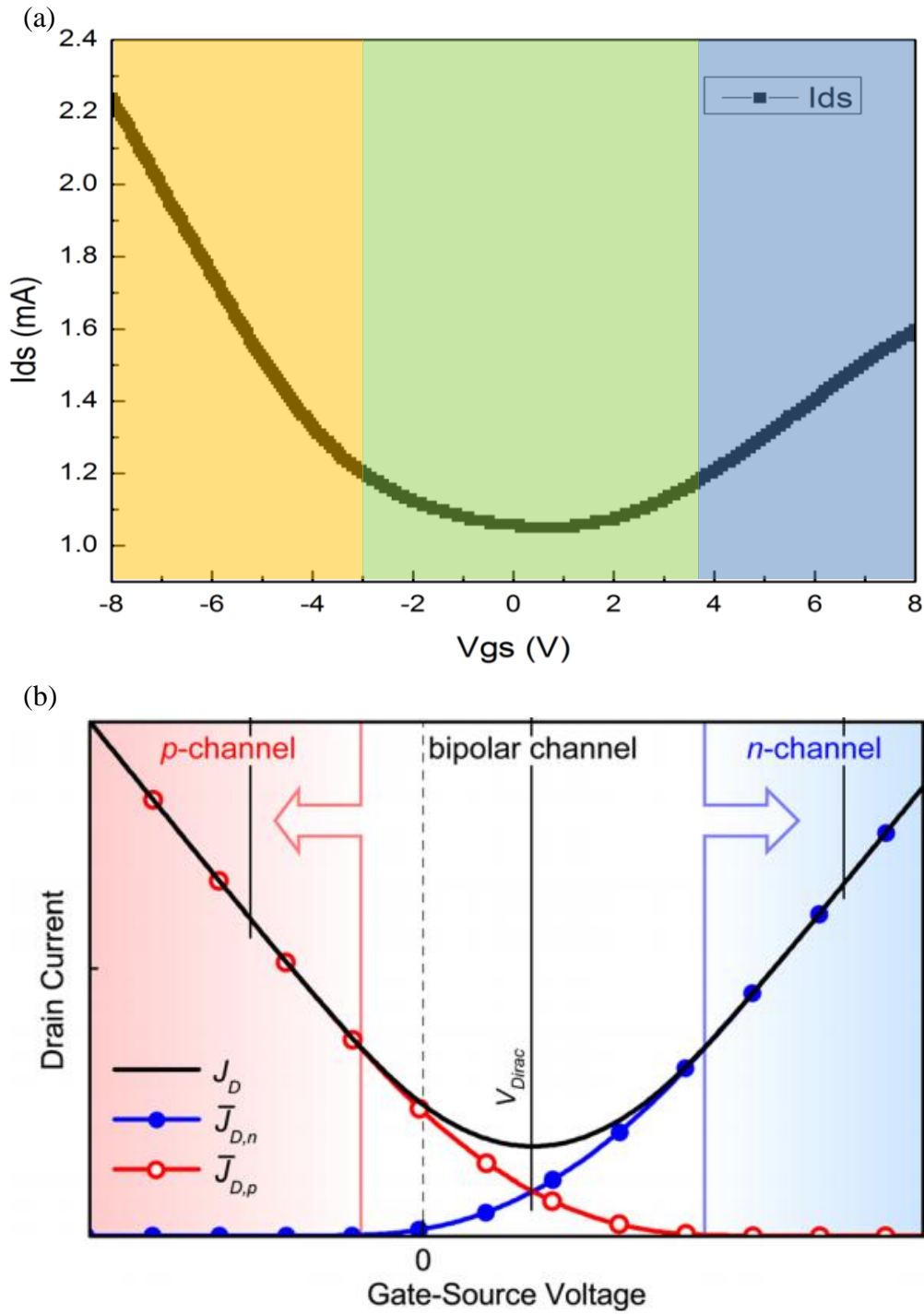
Figure 2.6 (a) Transfer curve of a CNTFET, with color differentiate p-type, ambipolar and n-type conduction region (b) schematic diagram of explanation for ambipolar conduction

majority carrier and this is the n-type conduction region. Since we are going to take advantage of the ambipolar region when use CNTFETs as frequency mixer, we are going to explain more about ambipolar conduction regime. For a single conduction channel transistor, either p-type or n-type, the point where the extended line of linear conduction region intercept with zero current axis is the threshold voltage. Threshold voltage is an important turning point where the transistor is considered stepping into linear regime. For an ambipolar transistor, it can be pictured such that p-channel having a threshold voltage larger than that of n-type threshold voltage, which means when gate voltage changing from negative to positive, n-channel opens before p-channel is closed and the region between n-/p-type threshold voltage can be viewed as ambipolar region. So ambipolar is not a proper of materials but rather a characteristic for certain device. In Figure 2.6 (b), the red line represents the current carried by holes and blue line represents the current carried by electron. It is clear that the ambipolar region is the result of overlapping the n-channel and p-channel current. Outside the ambipolar region, p-channel and n-channel do not influence with each other, and this is also the case for our CNTFETs.

Figure 2.7(a) shows the output curve in the p-type conduction region. V_{gs} changes from -10V to -2V. Still, we do not have sub-threshold in the output curve because of the existence of metallic nanotubes in the channel. Figure 2.7 (b) is the 2D view of the absolute value of current plotted vs. both V_{gs} and V_{ds}. The largest current in the channel is over 6.5mA, and no obvious saturation is observed. The high current carried by our CNTFETs can be attributed to two reasons: the first reason is the relatively high density of carbon nanotubes in the channel, and the top gate structure

32

also makes the transistor stay robust. The second reason is the Pd contacts, which

forms ohmic contacts with the carbon nanotubes. The contact resistance between

carbon nantobues and metal contacts is also low due to good wetting property of Pd.
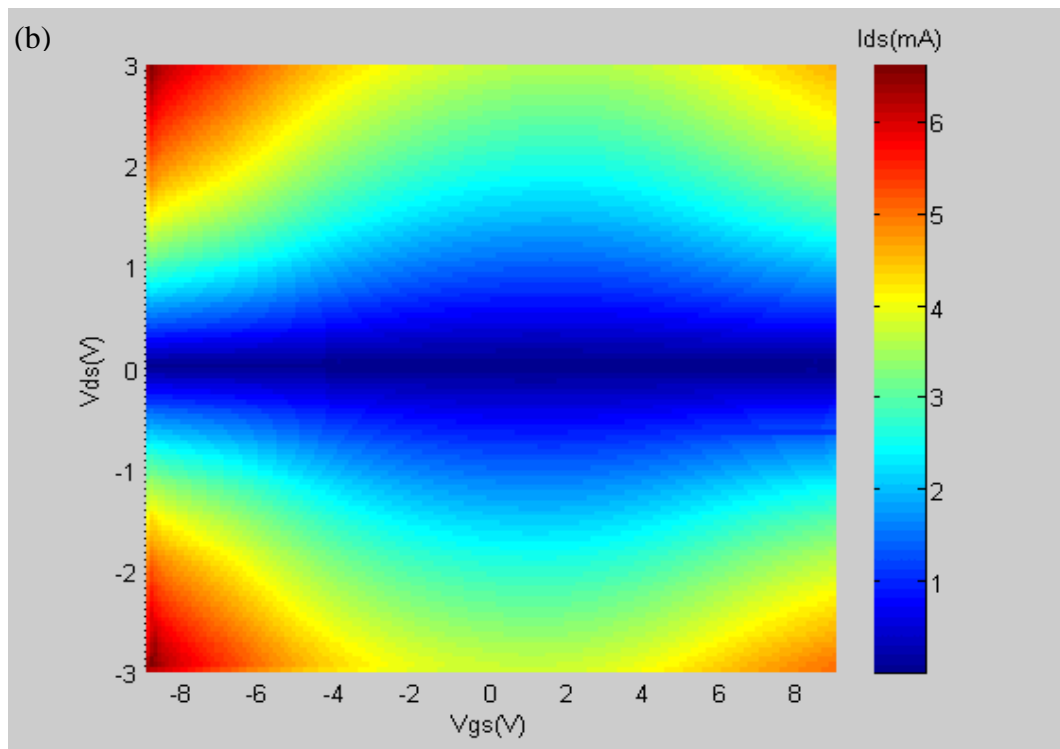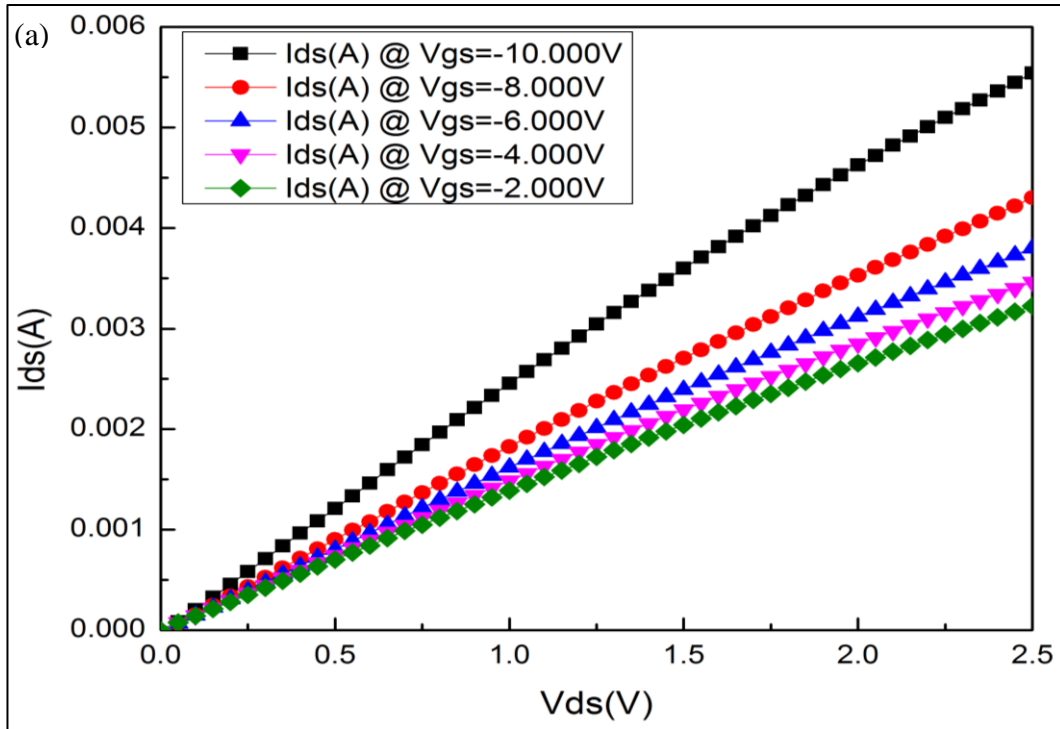
Figure 2.7 (a) Output curve of a CNTFET at different gate to source voltages (b) absolute value of current plotted vs. both V_{gs} and V_{ds} in a 2D view

Chapter 3: Harmonic generation from CNTFETs

3.1. Harmonic generation theory

First look at the Taylor expansion of drain current. The current between source and drain I_{ds} is a function of both V_{gs} and V_{ds}, which are the gate to source voltage and bias between source and drain. We can then write the equation

$$I_{ds} = f(V_{gs}, V_{ds})$$

Take this equation and do the Taylor expansion around a given voltage operation point with small AC signal (v_{gs}, v_{ds})

$$I_{ds} = f(V_{gs} + v_{gs}, V_{ds} + v_{ds})$$

$$= f(V_{gs}, V_{ds}) + \frac{\partial f}{\partial V_{gs}} \cdot v_{gs} + \frac{\partial f}{\partial V_{ds}} \cdot v_{ds} + \frac{\partial^2 f}{\partial V_{gs}^2} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial^2 f}{\partial V_{ds}^2} \cdot \frac{v_{ds}^2}{2!}$$

$$+ \frac{\partial^2 f}{\partial V_{gs} \partial V_{ds}} \cdot \frac{v_{gs} v_{ds}}{2!} + \cdots$$

Use the definition of transconductance and dynamic conductance here,

$$g_m = \frac{\partial I_{ds}}{\partial V_{gs}} \qquad G_d = \frac{\partial I_{ds}}{\partial V_{ds}}$$

We have

$$I_{ds} = f(V_{gs} + v_{gs}, V_{ds} + v_{ds})$$

$$= f(V_{gs}, V_{ds}) + g_m \cdot v_{gs} + G_d \cdot v_{ds} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial G_d}{\partial V_{ds}} \cdot \frac{v_{ds}^2}{2!}$$

$$+ g_m G_d \cdot \frac{v_{gs} v_{ds}}{2!} + \cdots$$

And furthermore, if we are only applying AC signal from the gate, all the derivatives related to V_{ds} can be omitted. And what we get is

$$I_{ds} = f(V_{gs}, V_{ds}) + g_m \cdot v_{gs} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial^2 g_m}{\partial V_{gs}^2} \cdot \frac{v_{gs}^3}{3!} + \cdots$$

Since the $f(V_{gs}, V_{ds})$ is the DC operation current of the transistor, let us define it as

I_{ds0}, and we have our Taylor expansion equation when the AC signal is applied

through the gate of our CNTFETs,

$$I_{ds} = I_{ds0} + g_m \cdot v_{gs} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial^2 g_m}{\partial V_{gs}^2} \cdot \frac{v_{gs}^3}{3!} + \cdots$$

The small signal AC current is then

$$i_{ds} = g_m \cdot v_{gs} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial^2 g_m}{\partial V_{gs}^2} \cdot \frac{v_{gs}^3}{3!} + \cdots$$

Assuming that the AC signal have a cosine wave form,

$$v_{gs} = A coswt$$

where A is the amplitude of the AC signal and w is the angular frequency. The square

and cube of v_{gs} are

$$v_{gs}^2 = (Acoswt)^2 = \frac{1}{2}A^2 + \frac{1}{2}A^2 cos2wt$$

$$v_{gs}^3 = (Acoswt)^3 = \frac{1}{4}A^3 cos3wt + \frac{3}{4}A^3 coswt$$

So as long as the derivatives before higher order terms are not 0, we will have double,

triple and multiple frequency harmonics generated. However, taking a close look at

the equation, besides the second order in the square equation, we also have a constant

with the amplitude of $\frac{1}{2}A^2$, and this corresponds to a DC current generated at by the

applied AC signal. Similar to the square equation, for the cube equation, we also have

additional $\frac{3}{4}coswt$ with the same frequency of fundamental signal. But since the

36

assumption here is v_{gs} is very small compared to the DC voltages, so the amplitude A

is also very small. So the main contribution to higher order harmonics is still from the

corresponding equation with same exponential order. The complete analysis of

harmonic generation from electronic devices needs the utilization of Volterra series.

However, since our main approach of relating the data to theoretical model is taking

numerical derivatives of raw data, which makes higher order contributions very

noisy, and for this reason, we will not use the full analysis of harmonic generation

including more higher order terms.

So based on the two assumptions we have here (i) only the nonlinearity of the

transconductance affects the harmonic generation since the RF signal is only applied

to the gate; (ii) the harmonics in the output power is mainly contributed from a

corresponding single exponential term in the Taylor expansion. We will investigate

the 1st harmonic generation from our CNTFETs and compare the transconductance

nonlinearity with the 1st order harmonic amplitude. The second order term in the AC

current is

$$\frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}{}^2}{2!} = \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{\left(\frac{1}{2}A^2 + \frac{1}{2}A^2 cos2wt\right)}{2!}$$

So the small current 1st order harmonic amplitude is

$$\frac{1}{4}A^2 \frac{\partial g_m}{\partial V_{gs}} \propto derivative\ of\ transconductance$$

Based on the above equation, we will set up single tone excitation measurement on

both CNTFETs and control devices to confirm the generation of harmonics, conduct

mapping measurement of 1st order harmonic amplitude and compare it to the

numerical derivative of transconductance to gate voltage.

3.2. Single tone excitation measurement

The experimental setup of single tone excitation measurement is shown in the

following block diagram
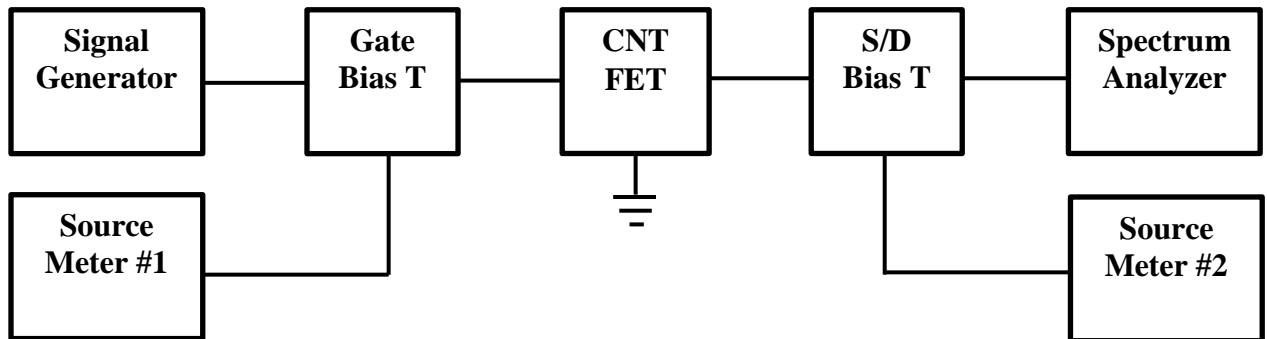
Figure 3.1 Experimental setup of single tone excitation measurement

The set-up components are listed below:

  Agilent MXG Analog Signal Generator (N5183A): 100 kHz ~ 40 GHz

  Agilent EXA Series Spectrum Analyzer (N9010A): 10 Hz ~ 44 GHz

  Agilent ENA Series Network Analyzer (E5071C): 300 kHz~20 GHz

  Keithley 2400 Source-Meters: DC

  Anritsu K250 Bias Tees: DC ~ 40 GHz

  Picoprobe 150 μm pitch microwave probes

For a typical single tone excitation setup, a 10GHz single with the power of

5dBm is generated from the signal generator, after coupling with the DC bias using

the bias T, transmitted into the CNTFET from the gate side. The DC bias will let the

device operate in different conduction type region (p-type, ambipolar or n-type). The

voltage between the source and drain is controlled using the source meter #2 through

the bias tee. The output signal containing both the fundamental signal and harmonics

generated by the CNTFETs will be passed into the spectrum analyzer, where the

relationship between harmonic amplitude and DC voltages will be measured.

Besides conducting measurements on CNTFETs, we also conduct the same

measurements on the control devices, which share identical device structure with

CNTFETs but do not have carbon nanotubes in the channel. By conducting such

control measurements, we can confirm that the harmonics generated are due to the

nanotubes in the channel instead of the device structure itself.

3.3. Single tone harmonic generation data

Based on the single tone excitation measurement setup mentioned above, we conduct

the measurement to observe the output spectrum when a 15dBm power is sent into the

gate. In Figure 3.2, the device under test is a $4\mu m$ channel length and $100\mu m$ channel

width CNTFET with gate length of 3.5 μm, and the input signal is 15dBm at 10GHz,

with the gate voltage and drain/source biases both as 0V. So the device is actually

working passively. However, even with the passive working condition, we observed

extra harmonic generation besides the fundamental signal at 10GHz, the 1st order

harmonic at 20GHz and the 2nd order harmonic at 30GHz is obvious on the output

spectrum.
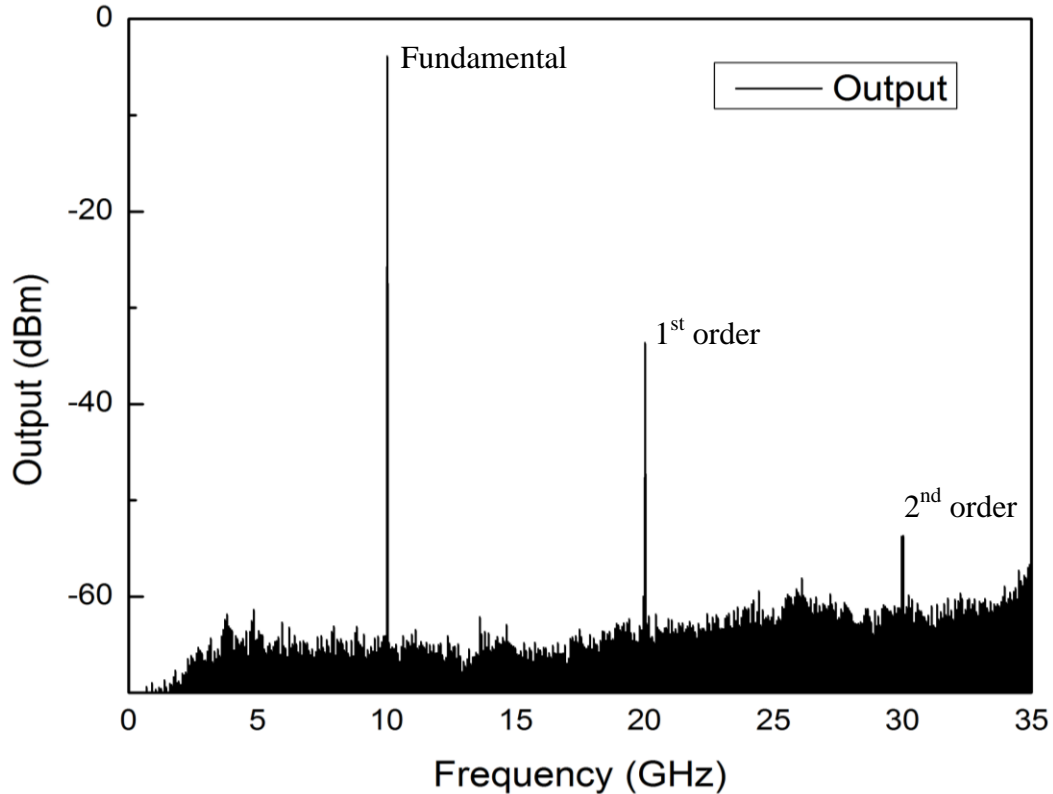
Figure 3.2 Output spectrum of single tone excitation measurement. The input power is 15dBm at 10GHz. Fundamental signal, 1st order and 2nd order harmonics are labeled.

The next measurement is the comparison between devices with nanotubes in the channel and without nanotubes in the channel while scanning the gate voltage of the device. For this measurement, the device is a $5\mu m$ channel length and $100\mu m$ channel width CNTFET with gate length of $5\ \mu m$, and the input power is 10dBm at 10GHz, with gate voltage from -8V to 8V. However, no source/drain bias is applied. The data is show in Figure 3. 3. We can extract two facts from the plot: (i) first, the extra harmonic generations at 20GHz and 30GHz are definitely due to the carbon nanotubes in the channel instead of from our blank device structure, since the output signals measured at 20GHz and 30 GHz from the control device are down to noise floor (ii) the second important fact is that the generated harmonic signal can be

modulated by applied gate voltage, even when the device is working passively (i.e. no source/drain bias applied). Also, the 1st and 2nd order harmonics are influenced by the gate voltage in different ways. This may either due to different nonlinear sources or same nonlinear source with different second and third order derivatives.
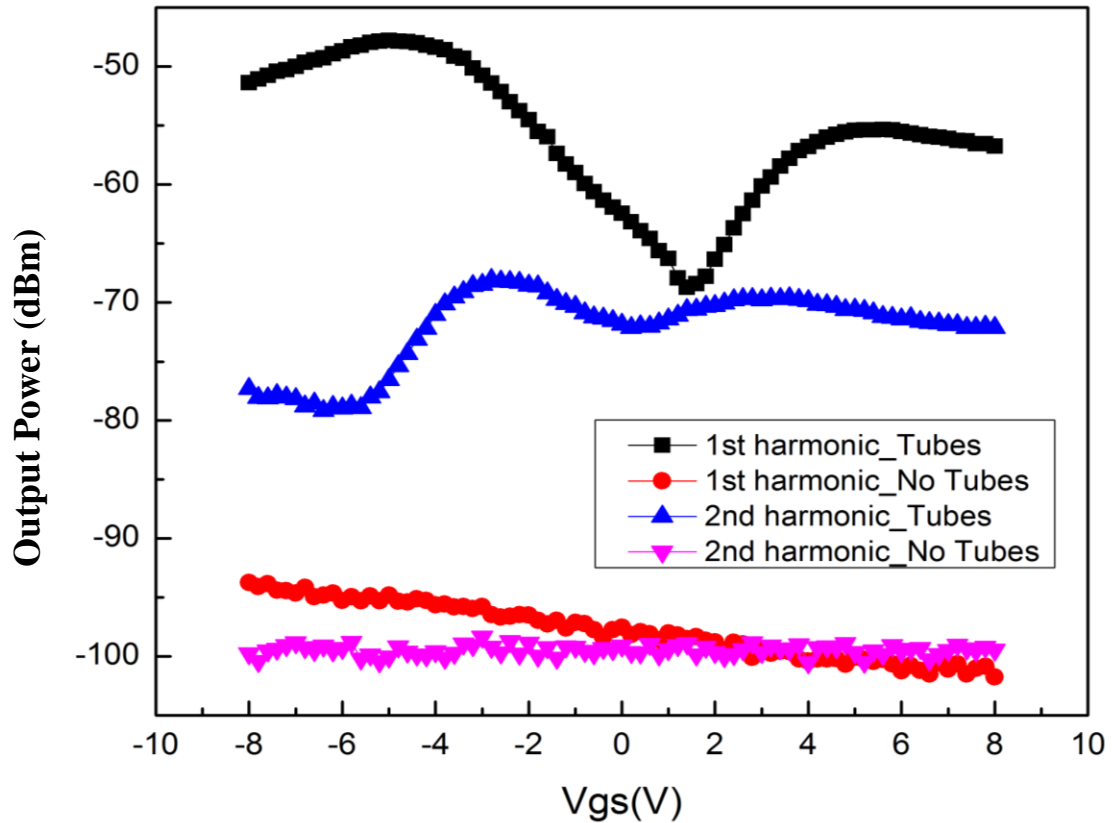
Figure 3.3 Output power measured at 20GHz and 30GHz when a 10GHz signal is applied to the CNTFET. Black and blue dots represent harmonics measured from device with nanotubes in the channel. Red and pink dots are data measured from control group without carbon nanotubes in the channel.

Note that we do not apply source/drain bias across our CNTFET channel in the previous measurements. In order to analyze how the active transistor behavior of the device influences the harmonic generation, we conduct a 2D (V_{gs}, V_{ds}) mapping test. In this measurement, the device under test is a $4\mu m$ channel length and $100\mu m$ channel width CNTFET with gate length of 3.5 μm, the input RF power is 5dBm at

41

10GHz. Figure 3.4 (a) shows the current mapping of the device in the voltage range of , Vgs from -8V to 8V and Vds from -2V to 2 V. Similar to previous DC tests, the device shows stronger p-type conduction with some extend of ambipolar conduction.

Figure 3.4 (b) is the mapping measurement results of 1st order harmonic amplitude at 20GHz. Since the way we do the mapping is scanning gate voltage for different source/drain biases, we get some shifting lines in the middle of the measurement. However, this will not influence the analysis of 1st harmonic generation. The figure shows that, the strongest 1st harmonic generation is within the gate voltage range of from -4V to 4V range. And in terms of source/drain bias range, it is small in around 0V bias line in the middle of the figure and gets stronger once it has a relatively large value, which means that the device is working as a transistor. Comparing the influence between V_{ds} and V_{gs} on the amplitude of 1st order harmonics, Vgs has a larger impact, and this is consistent with our experimental setup and theory, because the input signal is introduced through the gate, so nonlinear transconductance rather than nonlinear dynamic conductance will be the dominant role. Taking a close look at the mapping measurement data, the small 1st harmonic amplitude is from -8V to -4 V and from 4V to 8V, and within this range, the transfer curve of our CNTFETs is relatively linear, which means the second order term in the Taylor expansion or any higher order terms is small compared to the ambipolar region. We will further prove this when comparing the numerical derivatives of I_{ds} to the amplitude of harmonics.
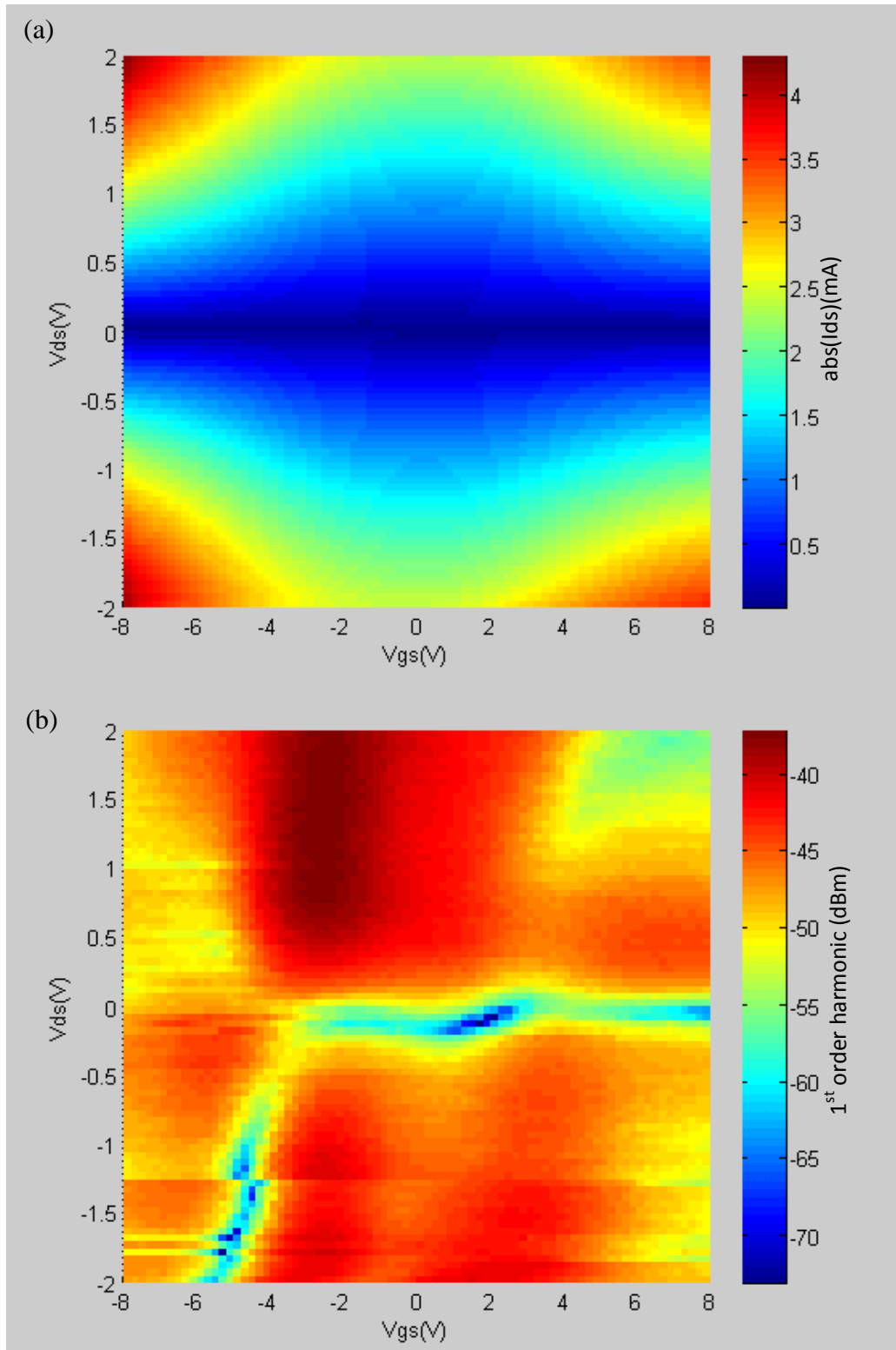
Figure 3.4 2D view of (a) absolute value of current and (b) 1st order harmonic amplitude plotted vs. both V_{gs} and V_{ds} in single tone excitation measurement

Based on the previous mentioned harmonic generation theory, the 1st order harmonic amplitude in the small AC signal is

$$\frac{1}{4} A^2 \frac{\partial g_m}{\partial V_{gs}} \propto derivative\ of\ transconductance$$

So we further calculate the numerical derivative of transconductance with respect to gate voltage and compare the value with the amplitude of 1st order harmonic. However, taking numerical derivative directly from raw data will produce significant noise, especially in our case where second derivative will be calculated based on raw data. So here we use the moving box method to take the numerical derivative. The method is described below. The measure current data I_{ds} is a matrix related to V_{gs} and V_{ds},

$$I_{ds} = \begin{pmatrix} I_{ds\ 1,1} & I_{ds\ 1,2} & I_{ds\ 1,3} & \cdots & I_{ds\ 1,m-1} & I_{ds\ 1,m} \\ I_{ds\ 2,1} & I_{ds\ 2,2} & I_{ds\ 2,3} & \cdots & I_{ds\ 2,m-1} & I_{ds\ 2,m} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ I_{ds\ n,1} & I_{ds\ n,2} & I_{ds\ n,3} & \cdots & I_{ds\ n,m-1} & I_{ds\ n,m} \end{pmatrix} \quad V_{ds} = \begin{pmatrix} V_{ds\ 1} \\ V_{ds\ 2} \\ \vdots \\ V_{ds\ n} \end{pmatrix}$$

$$V_{gs} = (V_{gs\ 1} \quad V_{gs\ 2} \quad V_{gs\ 3} \quad \cdots \quad V_{gs\ m-1} \quad V_{gs\ m})$$

For each $V_{ds\ k}$, I_{ds} will be an array with same length as V_{gs}. Take the first 5 elements in this array and linear fit, and assign the slope of the linear fit as the 1st derivative for element $I_{ds\ k,3}$. Take the $I_{ds\ k,2}$ to $I_{ds\ k,6}$, do the linear fit and assign the slope as the 1st derivative for element $I_{ds\ k,3}$. Keep doing this until we have the 1st derivative assigned for element $I_{ds\ k,m-2}$. Repeating this for every V_{ds} until we have all the rows covered. By doing this, we are considering 5 points instead of 2 points, which will lower the degree of noise to a great extent. We lose the derivative value for both first two and last two points, in comparison with only losing the first and last point in the array; this is the trade-off for lower noise.
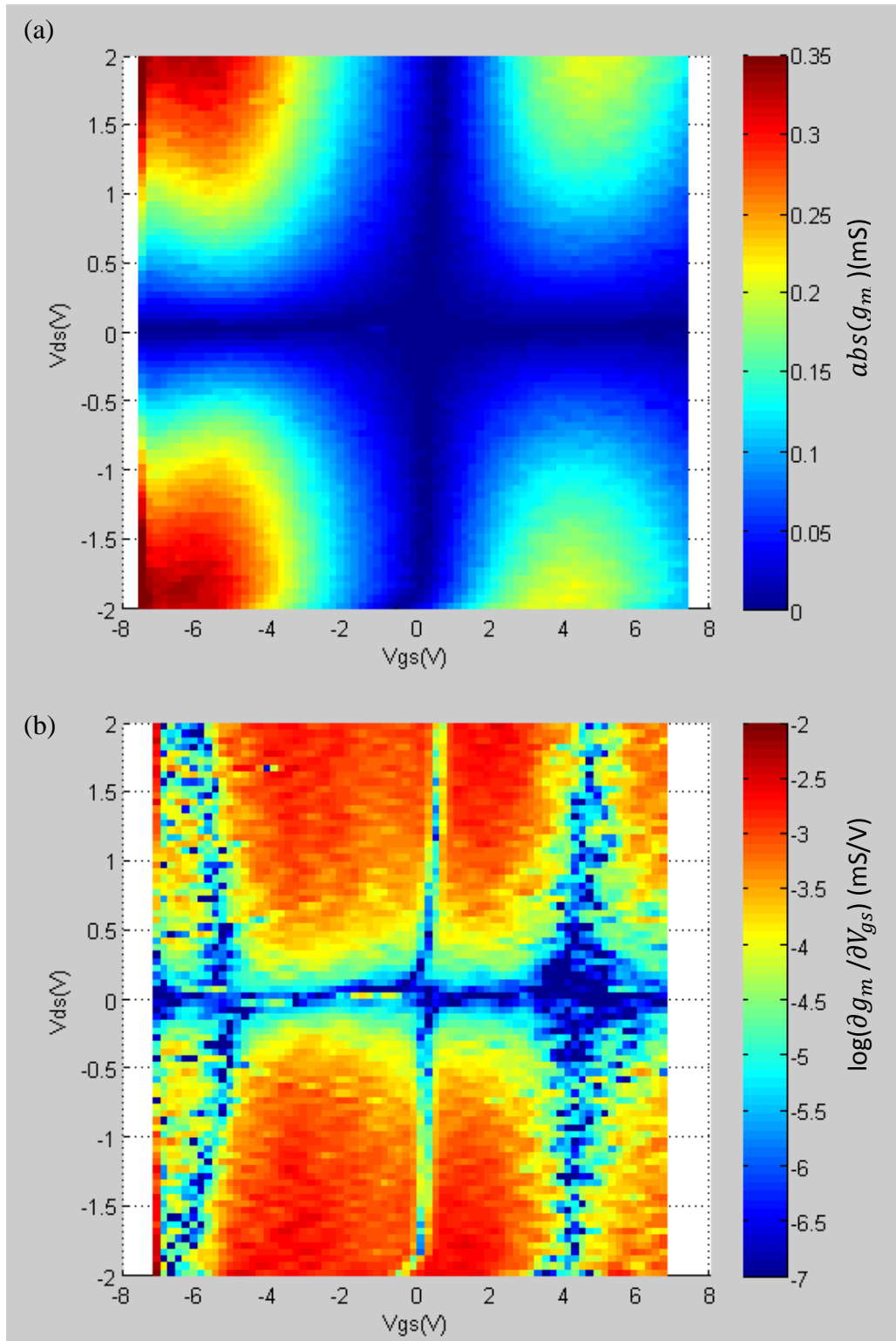
Figure 3.5 2D view of (a) transconductance and (b) logarithm of $\partial g_m / \partial V_{gs}$ plotted vs. both V_{gs} and V_{ds}

After conducting moving boxcar method to get the 1st derivative, we conduct moving boxcar method again to the 1st derivative matrix, and get the second derivative of current with respect to gate voltage, which is the derivative of transconductance with respect to gate voltage.

The numerical derivatives achieved by moving boxcar method are shown in Figure 3.5. In Figure 3.5 (a), the absolute value of transconductance is plotted vs. both V_{gs} and V_{ds}. Thanks to the moving boxcar method, the data shows smooth trend without noise spikes. The maximum of transconductance is in the p-type conduction region, with a value of ~0.35mS. Figure 3.5(b) shows the logarithm of $\partial g_m / \partial V_{gs}$ plotted vs. both V_{gs} and V_{ds}. The reason it is in logarithm scale is for the comparison with 1st harmonic amplitude, the unit of which is dBm. Even with moving boxcar method, we still have some extend of noise in (b), however, it is good enough for us to tell the trend. Compare Figure 3.4(b) with Figure 3.5(b), we notice that the larger $\partial g_m / \partial V_{gs}$ value corresponds to strong 1st order harmonic. This justifies our theory and suggests that the harmonic generation should be attributed to the nonlinearity of transconductance. This means that we can relate the different regions in DC transfer curve to the amplitude of harmonics.

Some differences are still seen. One possible reason is that higher order terms in the Taylor expansion also contributes to 1st order harmonics. However, based on the similarity between the two figures, 2nd order term in the Taylor expansion with coefficient of $\partial g_m / \partial V_{gs}$, is dominant in the generation of 1st order harmonic.

3.4. CNTFETs operated as frequency doubler

All the above single tone exciation measurements are conducted at several GHz regime, which is beyond the expected cut-off frequency of our CNTFETs. At lower frequency, the fundanmental output signal will be greatly modulated by the gate voltage. Based on the harmonic generation theory, the fundamental signal in the output spectrum is $g_m \cdot v_{gs}$. From Figure 3.5 (a), the minimum of transconductance lies between 0V to 1V gate voltage, which is around the minimum current point. If the CNTFET is working around this gate voltage point, the fundanmental output signal will be greatly suppressed and even smaller than the 1st order harmonic, which makes the signal with double frequency dominates the spectrum[65]. Thus an effective frequency doubler can be built with CNTFET[66].

Figure 3.6 shows the output spectrum of a CNTFET operated as frequency doubler. The input signal has frequency of 5MHz and amplitude of 10dBm. The transistor is biased to $V_{gs} = 0.5V$ and $V_{ds} = 1.5V$. As is expected, the fundamental signal is greatly suppressed, but the 1st order harmonic is within its strong V_{gs} regime. Thus, we observe a doubled frequency signal that is about 15dBm larger than the fundamental signal. An effective frequency doubler normally requires the double frequency signal to be at least 10dBm larger than the fundamental output signal. And our CNTFET certainly meets this requirement. The capability of CNTFET to be operated as frequency doubler is due to the shape of transfer curve in the ambipolar region. Near the minimum current point, the transfer curve is near the shape of a parabola, whose 1st derivative around minimum current point is 0 and has a large second derivative. The reason we still have considerable fundamental signal in the output spectrum is because the shape of
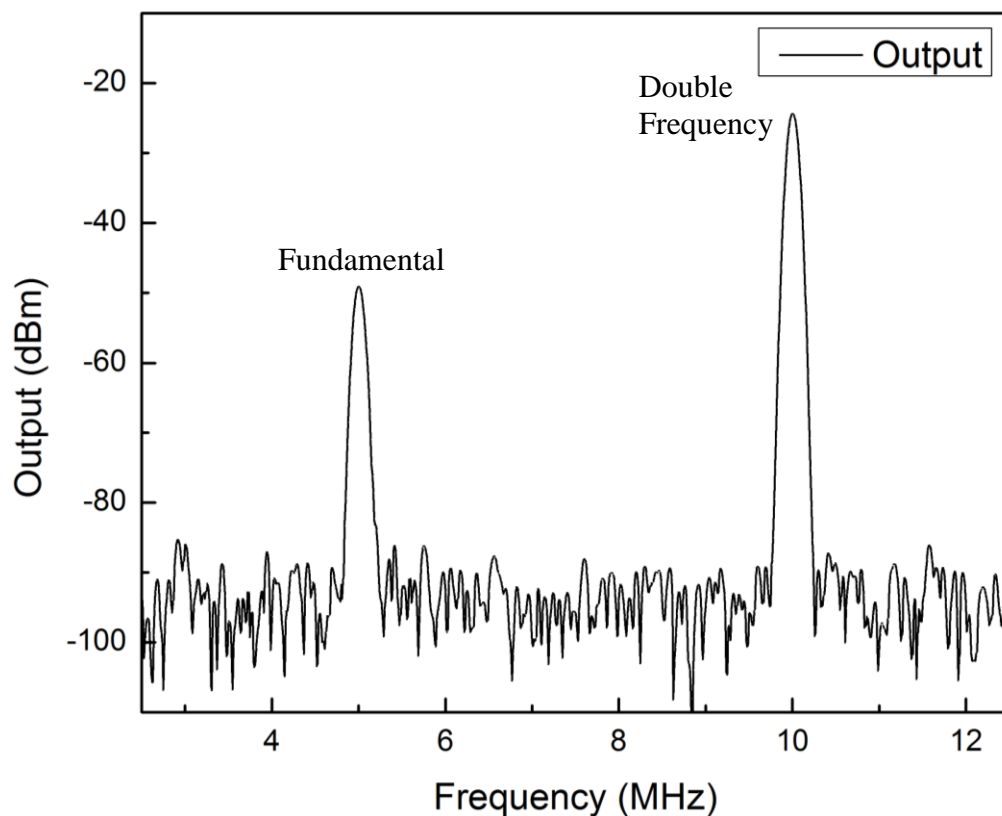
Figure 3.6 Output spectrum of CNTFET operated as frequency doubler. The input signal is at 5MHz with the amplitude of 10dBm.

the transfer curve is not perfect parabola, which lead to a non-zero 1st derivative at the

operation point. The 15dBm difference between double frequency signal and

fundamental output signal is sufficient.

Chapter 4: Elementary model of nonlinearity in CNTFETs

4.1. Drude Model

Drude model is a classical model to explain the transport properties of electrons in

materials. And in the model, it introduced important concepts such as mean free path.

The simplest analysis of the Drude model assumes that electric field **E** is both

uniform and constant. It assumes that electrons collide with scattering centers and

regain momentum after colliding with another scattering center, and τ is the average

time between collisions. Within τ, the average momentum electron gains is

$$\Delta\langle \boldsymbol{p}\rangle = q\boldsymbol{E}\tau$$

During its last collision, this electron will have been just as likely to have bounced

forward as backward, so all prior contributions to the electron's momentum may be

ignored, resulting in the expression

$$\langle \boldsymbol{p}\rangle = q\boldsymbol{E}\tau$$

After that, substitute the classical kinetic equation

$$\langle \boldsymbol{p}\rangle = m\langle \boldsymbol{v}\rangle$$

and current density equation

$$\boldsymbol{j} = nq\langle \boldsymbol{v}\rangle$$

This results in the formation of Ohm's law equation

$$\boldsymbol{j} = \left(\frac{nq^2\tau}{m}\right)\boldsymbol{E}$$

As is mentioned, even though this is a very simple case of Drude model, it opens

up questions when electron transport happens in extremely small objects, where it is

even smaller than the mean free path of electron. As is shown in Figure 4.1, there are

two important lengths when considering whether the carrier is following ballistic or

diffusive transport characteristics. The first length is mean free path of the electron

(holes), which is defined as distance an electron travels until its initial momentum is destroyed, named as L_e. The second length is the channel length between contacts, named as L.

(i) When $L > L_e$, carriers display diffusive transport characteristics, which is shown in Figure 4.1(a). Scattering is the main contribution to the resistance of the channel.

(ii) When $\lambda_F < L < L_e$, carriers display ballistic transport characteristics, which is shown in Figure 4.1(b). Contacts are the main contribution to the resistance of the channel.
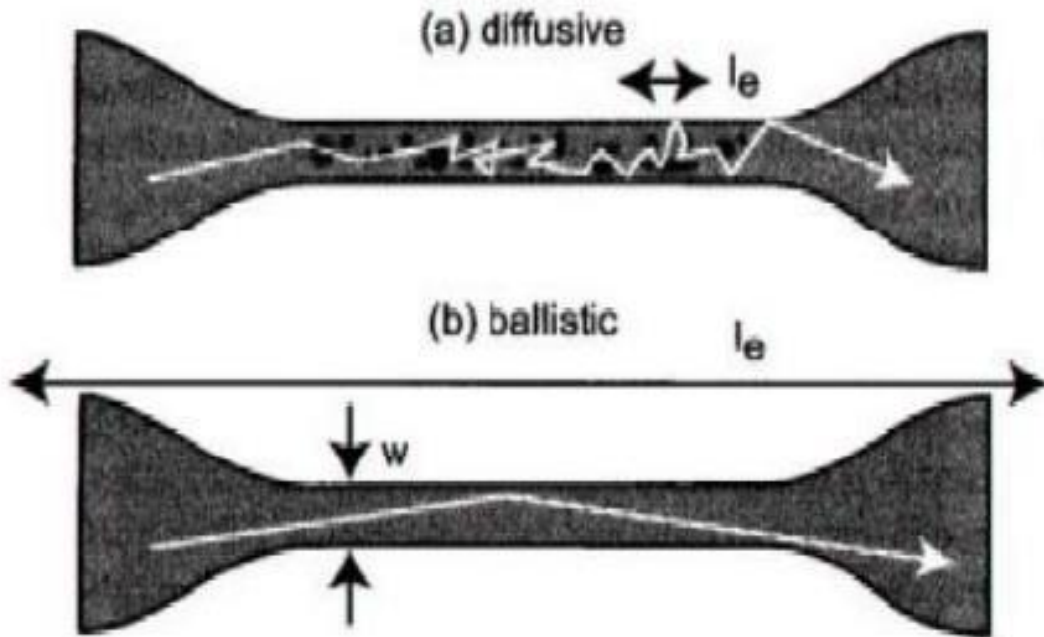
Figure 4.1 Schematic drawing of diffusive and ballistic transport characteristics. (a) shows diffusive transport when $L > L_e$, and (b) shows ballistic transport when $L > L_e$

4.2. Ballistic electron transport

Generally, ballistic transport will happen when the length of the channel is smaller than the mean scattering length. Consider a conductor connected to two electrodes, as is shown in Figure 4.2, the length of the conductor is L. E_{F1} and E_{F2} are the Fermi level of electrode #1 and #2, which are kept as constants. Due to the quantization of wave vectors k perpendicular to L direction, there are several subbands that will contribute to the total current, and M is the number of subbands, also called number of channels. The velocity of electron is

$$v = \frac{1}{\hbar}\left(\frac{\partial E}{\partial k}\right)$$

where ℏ is the reduced Plank constant. Current can be written as

$$I = q/t_t$$

$$t_t = L/v$$

t_t is the carrier transit time and q the charge of carrier. Based on this, the total current is

$$I = \frac{q}{L}\sum_{M,k}\frac{1}{\hbar}\frac{\partial E}{\partial k}[f(E - E_{F1}) - f(E - E_{F2})]$$

$$= \frac{q}{L} \times 2 \times \frac{L}{2\pi}\sum_{M}\int dk \frac{1}{\hbar}\frac{\partial E}{\partial k}[f(E - E_{F1}) - f(E - E_{F2})]$$

$$= \frac{2q}{h}\int dE[f(E - E_{F1}) - f(E - E_{F2})]M$$

$$\approx \frac{2q^2}{h}\frac{[E_{F1} - E_{F2}]}{q}M$$

$$= \frac{2q^2}{h}VM$$

where f stands for Fermi function. In the conversion from sum of k to integral, the 2

comes from the spin degeneracy and $\frac{L}{2\pi}$ is the inverse of the level spacing. $\frac{[E_{F1}-E_{F2}]}{q}$ is

substitute into V, the voltage between two contacts. In the case of coherent transport,

where the phase and amplitude of wave function at electrode #2 can be achieved from

electrode #1. A new parameter is introduced, which is the transmission probability, T.

It stands for the transmission probability for a channel extending from #1 to #2

electrode. Thus the conductance of the channel is

$$G = \frac{I}{V}T = \frac{2q^2}{h}MT$$

the equation is called Landauer formula. It illustrates the origin of resistance even

when the material in the channel is shorter than the mean free path of the carrier.

Ballistic CNTFETs have been fabricated and studied[55][67][68], and show superior

performance both from experimental work and theoretical projection[69][70][71].

Our CNTFETs have longer channel than the mean free path of carriers, so ballistic

transport is not applicable. But we can still use the 1D conductance equation.
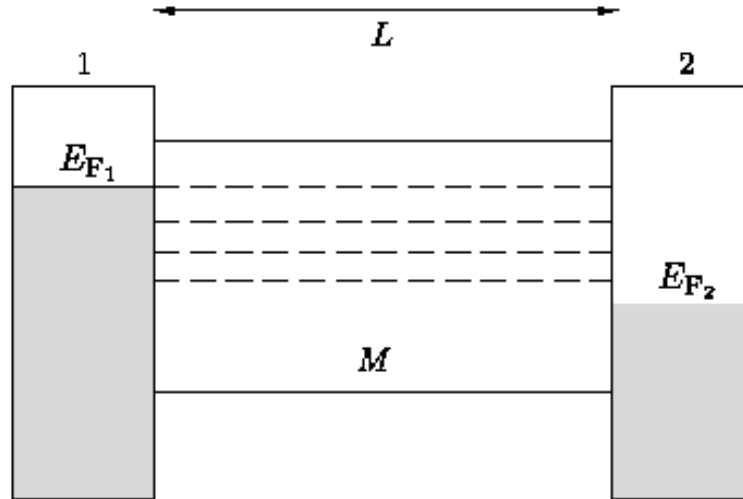
Figure 4.2 Schematic drawing of ballistic transport in a short channel. The length of the channel is L. #1 and #2 contacts have Fermi energy of E_{F1} and E_{F2}. M represents the number of sub-bands in the conducting channel.

4.3. Inherent linearity in CNTFETs

There are few studies on the linearity of CNTFETs[72][73], and one of them is a simplified calculation of current in the channel following two assumptions: (1) the contacts between carbon nanotubes and the metal contacts are Ohmic contacts (2) the electronic transport in the carbon nanotube is ballistic[56]. Among these two assumptions, Ohmic contacts can be achieved using Pd as contact metal. By using e-beam lithography to define the channel length of CNTFET, ballistic (or quasi-ballistic) transport can be achieved. Here we are going to examine the linearity of CNTFET following their path.

Since we assume ballistic transport, the schematic model is still Figure 4.2. However, we now name #1 and #2 electrodes as source and drain. Define the electrostatic potential of source, drain and carbon nanotubes as $V_s, V_d,$ and V_{cnt}, and the gate capacitance as C_g, the gate to source voltage as V_g. Laudauer formula

indicates that it requires energy for electron to enter the channel. And we name the energy required to enter the channel from source as channel potential μ. We now have the equation

$$-eV_s = -eV_{cnt} + \mu$$

The applied gate voltage will cause a linear distribution of charge in the channel

$$\lambda = C_g(V_{cnt} - V_g) = \frac{C_g}{e}(\mu - eV_g)$$

where λ is the linear charge distribution induced by applied gate voltage. On the other hand, we also have

$$\lambda = -e\int dED(E)f(E-\mu)$$

where D(E) is the density states of the lowest subband, $f(E)$ is the Fermi function and define $f(E)^{-1} \equiv 1 + \exp(\beta E)$, where β is 1/kT,

$$D(E) = \frac{4}{hv_F}\frac{E}{\sqrt{E^2 - (\frac{E_g}{2})^2}}$$

where E_g is the bandgap of carbon nanotube, v_F is the Fermi velocity of carbon nanotube. Notice that here we need E larger than half of E_g. Combine all the above equations together and we can get

$$\frac{hv_F}{4e^2}C_g(eV_g - \mu) = \int_{\frac{E_g}{2}}^{\infty}\frac{dE}{1 + \exp[\beta(E-\mu)]}\frac{E}{\sqrt{E^2 - (\frac{E_g}{2})^2}}$$

Based on this equation, we can solve the potential in the carbon nanotube channel μ. Use the definition of quantum capacitance,

$$C_Q = \frac{4e^2}{hv_F}$$

We have

$$\frac{C_g}{C_Q}(eV_g - \mu) = \int_{\frac{E_g}{2}}^{\infty} \frac{dE}{1 + \exp[\beta(E - \mu)]} \frac{E}{\sqrt{E^2 - (\frac{E_g}{2})^2}}$$

And further utilize the 1D electron transport relationship

$$v = \frac{\partial E}{\hbar \partial k} = \frac{L}{h} D(E)^{-1}$$

We can write down the current I as a function of μ

$$I(\mu) = \frac{4e}{h} \int_{\frac{E_g}{2}}^{\infty} dE \cdot T \cdot f(E - \mu)$$

$$= \frac{4e}{h\beta} \ln(1 + \exp[\beta(E - \mu)])$$

T is the transimission probability and here we can assume it is one. The analysis

remains valid as long as T is independent of energy. Furthermore, if we consider that

we have very high-k materials as dielectric gate oxide, C_g can be greatly larger than

C_Q, and the solution to μ is limited to $\mu \approx eV_g$, this makes the current expressed as

$$I = \frac{4e}{h\beta} \ln(1 + \exp\left[\beta(eV_g - \frac{E_g}{2})\right])$$

Notice that this is logarithm of an exponential. It will become linear when very small

gate voltage is applied, β will be cancelled. So we got the inherent linearity of

CNTFETs

$$I = \frac{4e^2}{h}(V_g - V_T)$$

where V_T is the threshold voltage and $V_T = \frac{E_g}{2}/e$.

So far, we get the inherent linearity of channel current's dependence on gate to source voltage. However, there is discrepancy within the assumption that C_g can be greatly larger than C_Q. Even with HfO_2 as gate dielectric layer, and using atomic layer deposition to form very thin oxide films, we cannot get such a large gate capacitance. So the linearity will be deteriorated due to this, and in the most practical case, C_g is only a fraction of C_Q, which makes to total capacitance depends on C_g more.

Besides the capacitance assumption, there is another reason that the inherent linearity model cannot fit perfectly to our case. One of the assumption is carriers in carbon nanotube display ballistic transport characteristics. In our device, the channel length is from $2\mu m$ to $10 \mu m$. The mean free path of electrons in carbon nanotubes is only several hundred nanometers. So scattering is crucial to the transport of carriers in our CNTFET device. We will modify the ballistic model and combine it with the 1D Drude model to calculate the current in the channel.

4.4. Modeling the current transport and nonlinearity in CNTFETs

We start with the conductance of 1D diffusive transport with 4 channels[74], since we have relatively long carbon nanotubes in comparison with the mean free path, and the conductance G is

$$G = \frac{4e^2}{h}\frac{L_m}{L_c}$$

where L_m is the mean free path length and L_c is the channel length of the CNTFET. The conductivity σ is then a function of the length x along the tube, define it as $\sigma(x)$.

$$\sigma(x) = \frac{4e^2}{h}l(x)$$

56

where $l(x)$ is the mean free path at x. For 1D conductor, the scattering rate is proportional to density of states[75][76]. So we write

$$l(x) = l_0 \left(\frac{v(x)}{v_0}\right)^2$$

where $v(x)$ is the Fermi velocity along the channel and v_0 is the Fermi velocity at high energy. l_0 is the mean free path at high energy, $l_0 = v_0 \tau_0$. τ_0^{-1} is the total scattering rate and

$$\tau_0^{-1} = \alpha \frac{T}{d}$$

where T is temperature and d is the diameter of the carbon nanotube. The value of α is experimentally measured, and have the value of about $12 mK^{-1}s^{-1}$[77]. Using the relativistic band structure of carbon nanotube[78], which is

$$E = \pm \sqrt{(m^* v_0^2)^2 + (\hbar k v_0)^2}$$

we can have the relation between v(x) and k,

$$\left(\frac{v(x)}{v_0}\right)^2 = \frac{\left(\frac{\hbar k}{m^* v_0}\right)^2}{1 + \left(\frac{\hbar k}{m^* v_0}\right)^2}$$

Here, m^* is the effective mass of the carrier and we can also relate k with the applied gate voltage

$$k(x) = \frac{\pi}{4} \frac{C_g}{e} (V_g - V(x))$$

where C_g is the gate capacitance and V_g is the gate voltage applied. Distinguish this V_g with the one in the inherent linearity of CNTFETs, since this V_g is the gate voltage to

ground, and source is grounded. And then we can finally write down the current for

one carbon nanotube in the channel,

$$I_{ds} = \frac{1}{L_c} \int_0^{L_c} \sigma(x)E(x)dx = \frac{1}{L_c} \int_0^{V_{ds}} \sigma(V)dV$$

Our model of calculating the current in the channel is based on this equation. And

the method we use to characterize the nonlinearity is taking derivatives of current

with respect to gate voltage. Also, we treat the capacitance of arrays of carbon

nanotubes as sum individual carbon nanotubes. This is because the distance between

our nanotubes in the channel is about $0.3\mu m$ and this is very large compared to the

diameter of carbon nanotubes and the thickness of dielectric layer (60nm).

Three other points deserve attention:

 (1) the number of carbon nanotubes in the channel is ~2 per micron, which

makes a $100\mu m$ double-channel device have 400 nanotubes responsible for the total

current, it is very hard to characterize or model the electronic transport for each

carbon nanotubes and put them together. So we simplify the problem by assuming all

semiconducting nanotubes having the same band structure, and this will be reflected

on the value of effective mass m^* . By treating the m^* as a fitting parameter to get the

most alike current transport mapping with the measurement, we are averaging the

influence of different chirality and band structure. Another averaging is on the

diameter of carbon nanotubes. It requires a lot of work to accurately measure the

diameter of the carbon nanotubes and treat them separately. There is research work

studying the averaging effect of nanotube diameters and statistics effect on the

performance of CNTFETs[79][80]. In order to model the group behavior of carbon

nanotubes in the channel, a certain diameter is assigned and values around 1.5nm is viable based on the catalyst layer thickness and CVD growth conditions.

(2) Our CNTFETs are built based on CVD growth, which gives about one third of the nanotubes as metallic. But metallic nanotubes will not display gate control, nor will they be responsible for nonlinearity of the CNTFETs. So we do not consider the current from metallic nanotubes. The ignorance of metallic nanotubes will lead to difference between measured DC current and the modeled current, but will not influence the study of nonlinearity of CNTFETs.

(3) Quantum capacitance is not taken into consideration. This is because that, based on the structure of our device (60~70nm Al_2O_3 as dielectric layer, nanotubes have a diameter between 1nm to 2nm), the gate capacitance between carbon nanotubes is small compare to quantum capacitance of carbon nanotubes. A comparison between with/without quantum capacitance will show that there is basically no difference for our CNTFETs.

4.5. Modeling data

For all the modeling data, the calculation is only based on a single semiconducting carbon nanotube. The reason is that, the number of carbon nanotubes in the channel is only an approximation, so multiplication of the total number of nanotubes with the single nanotube characteristics will only add another fitting parameter. Based on our assumption that individual carbon nanotubes are not interacting with each other due to the large distance between tubes in comparison with the diameter, the trend displayed by an individual nanotube can represent the

overall behavior of CNTFET. And this is especially true when we plot our modeling

current data with respect to both V_{ds} and V_{gs}. The difference in quantity in DC

characteristic also does not affect the trend when comparing derivatives of current

with respect to V_{gs} and the measured harmonic amplitude.

Figure 4.3 shows the DC characteristics for an individual semiconducting nanotube

FET. The length of the carbon nanotube is set to 4μm, same as in the measurement

below. The effective mass m^* is set to $0.75m_e$ and the diameter is 1.5nm. v_0, Fermi

velocity at high energy is set to $8 \times 10^7 cm/s$. The dielectric is Al_2O_3 with the

thickness of 70nm, the relative dielectric constant is set to 9. In (a), it shows the

output curve of the single tube at different gate to source voltages. From -8V to -2V

with 2V as step, the current decreases accordingly, as is expected when the CNTFET

is working in the p-type conduction region. The upper limit of V_{ds} is set to 2V, the

same as in measurement. The largest current is at $V_{ds} = 2V$ and $V_{gs} = -8V$, with the

value of ~17μA. The output curve mainly stays within the linear region of transistor

operation, and this is also expected since the channel length is 4μm, so the maximum

electric field is only 0.5V/ μm, which is not strong enough to cause saturation in the

CNTFET. Figure 4.3 (b) shows the transfer curve of the individual carbon nanotube

FET, with a constant $V_{ds} = 1.6V$, and gate to source voltage sweeping from -8V to

8V. The transfer curve also shows three different conduction regions, which are p-

type region, ambipolar region and the n-type region. Note that the minimum current is

very close but not zero, and this is because of the small bandgap in our modeled

individual nanotube. The threshold voltage of p-type conduction is larger than that of

n-type conduction, so that within the ambipolar region, the minimum current point the

sum of subthreshold current of both p-type and n-tpye region. In terms of linearity, the transfer curve show good linearity at moderate bias voltage outside ambipolar reigion. However, at large gate to source voltage, for both p-type and n-type conduction regions, the transfer curve shows some extend of saturation. One possible reason for this saturation is the assumption that only one subband contribute to the conduction. In our calculation, we only considered the first subband, but it is very likely that at larger gate bias, another or even multiple subbands will contribute to the total current. Including another subband will make the slope at larger gate bias larger, and instead of saturation, it may keep displaying good linearity.
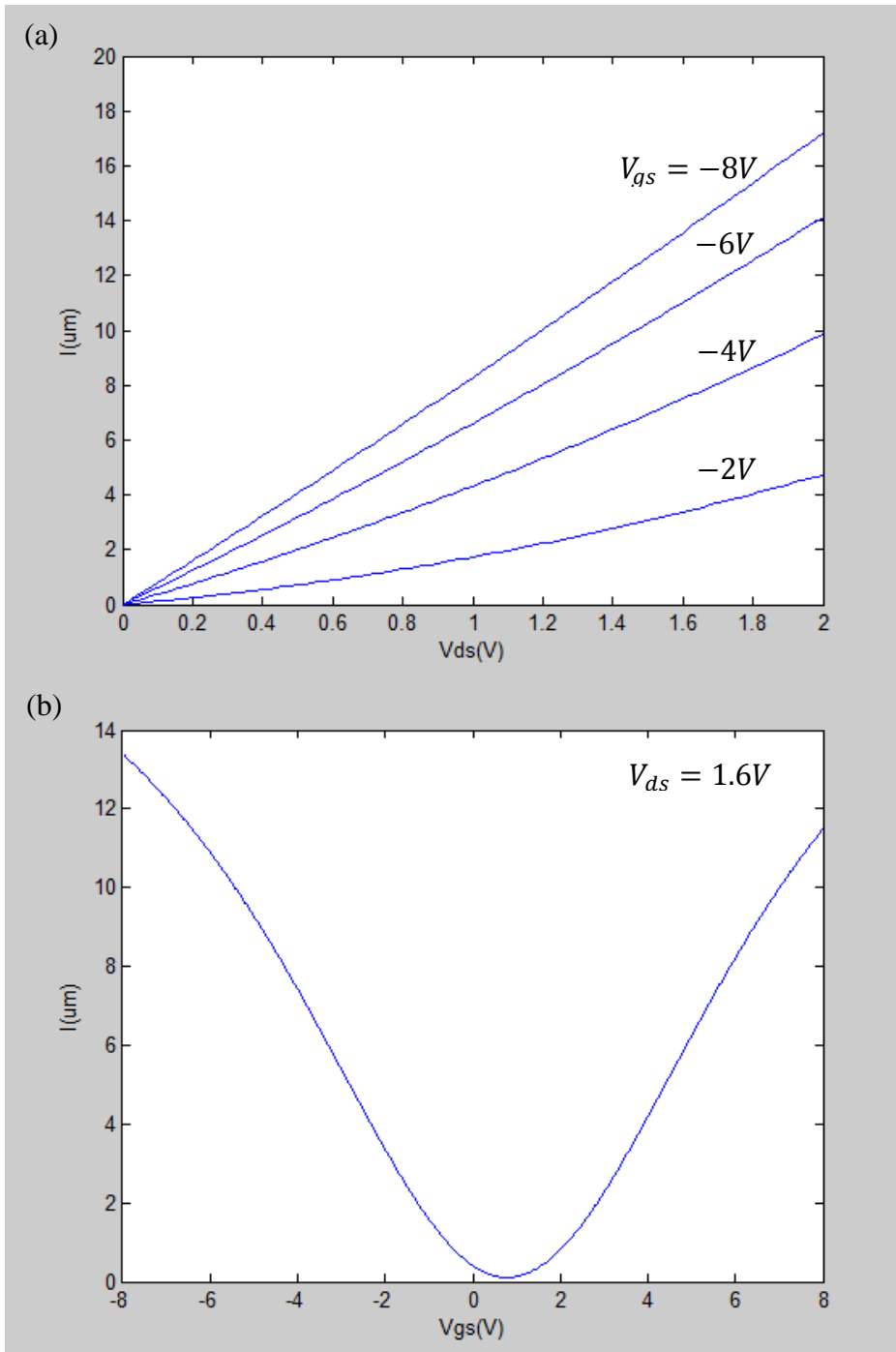
Figure 4.3 DC characteristics of a single semiconducting nanotube calculated from model. The effective mass m^* is set to $0.75m_e$ and the diameter is 1.5nm. (a) shows the output curve with V_{gs} changing from -8V to -2V (b) shows the transfer curve at $V_{ds} = 1.6V$.

In order to compare the overall current vs. voltage relationship, we calculate

the current at different V_{ds} and V_{gs}, and plot the data in a 2D mapping manner. Figure

4.4 (a) shows the calculated current vs. V_{ds} and V_{gs} based on the model from an

individual carbon nanotube FET. In the color mapping, red stands for high value of

current, and blue for low. The information we can extract from the map is that the

individual nanotube FET have a stronger p-type conduction region than n-type. The

minimum current point with respect of V_{gs} is around 1V. In Figure 4.4 (b), it is the

measured current plotted vs. V_{ds} and V_{gs}. This is a CNTFET with a channel length of

$4\mu m$, a channel width of $100\mu m$ and a gate length of $3.5\mu m$. The dielectric is Al_2O_3

with the thickness of 70nm. At first glance, the most significant difference between

(b) and (a) besides the scale is around $V_{gs} = 1V$ region. In Figure (b), around 1V gate

voltage, the current increases as the source/drain bias getting larger, which is not the

case in Figure (a). This is due to the existence of metallic carbon nanotubes in the

CNTFET under test. Around one third of the total nanotubes are metallic, and even at

the minimum current point, where the current should not, or only slightly on the

source/drain bias for semiconducting nanotubes, metallic nanotubes will carry current

increasing linearly with the source/drain bias. We can do a simple estimation of the

influence of metallic nanotubes, the largest current in (b) is 4.3mA, and if one third of

them is from metallic nanotubes, the current carried by semiconducting nanotubes is

4.3mA*2/3=2.87mA. The individual semiconducting nanotube in the model carries

$17\mu A$ current as maximum. So the approximate number of semiconducting nanotubes

in the channel is 2.87mA/$17\mu A$=~170. And this is a plausible number of

semiconducting nanotubes in the channel based on the density of carbon nanotubes.

Besides the difference around minimum current point, the model can predict the trend of the relationship between current and voltage when compared to the measured data.

The next step is to compare the nonlinearity in the model with what the amplitude of harmonics in the single tone excitation measurement. First, let us recall our derivation of harmonic generation. The small signal AC current from output signal is

$$i_{ds} = g_m \cdot v_{gs} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial^2 g_m}{\partial V_{gs}^2} \cdot \frac{v_{gs}^3}{3!} + \cdots$$

where g_m is the transconductance and is defined as

$$g_m = \frac{\partial I_{ds}}{\partial V_{gs}}$$

Assuming that the AC signal have a cosine wave form,

$$v_{gs} = A \cos wt$$

where A is the amplitude of the AC signal and w is the angular frequency. The square of v_{gs} is

$$v_{gs}^2 = (A \cos wt)^2 = \frac{1}{2} A^2 + \frac{1}{2} A^2 \cos 2wt$$

The second order term in the AC current is

$$\frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} = \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{\left(\frac{1}{2} A^2 + \frac{1}{2} A^2 \cos 2wt\right)}{2!}$$

So the 1st order harmonic amplitude is

$$\frac{1}{4} A^2 \frac{\partial g_m}{\partial V_{gs}} \propto derivative\ of\ transconductance$$
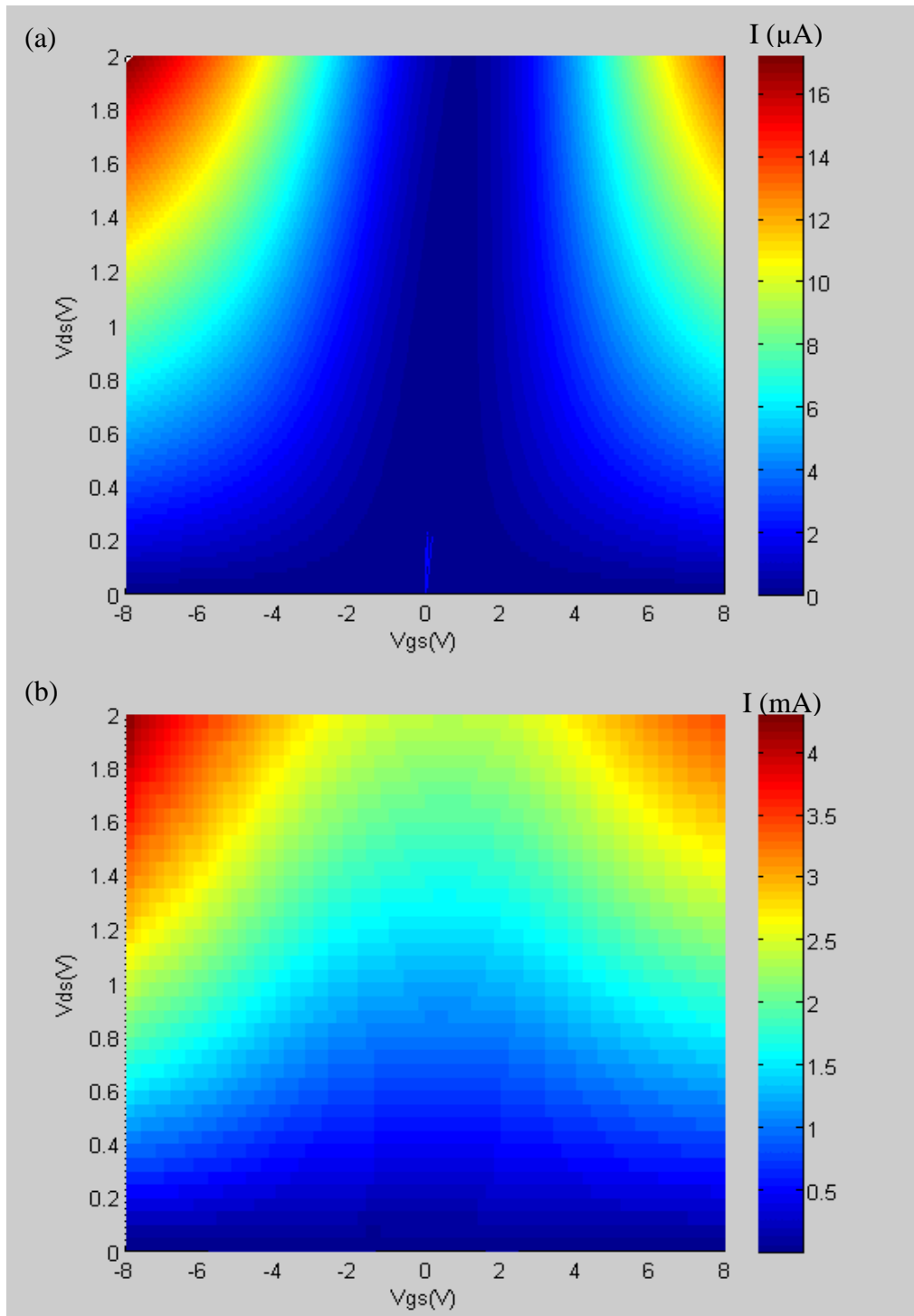
Figure 4.4 DC current mapping vs. V_{ds} and V_{gs} from (a) model and (b) measurement. In (a), the model still only calculate the current in a single carbon nanotube. In (b), the CNTFET under test has both metallic and semiconducting nanotubes in the channel

And the comparison we are going to make is between $\frac{\partial g_m}{\partial V_{gs}}$ and the amplitude of 1st

order harmonic amplitude. Use the definition of transconductance

$$\frac{\partial g_m}{\partial V_{gs}} = \left. d^2 I_{ds} \middle/ dV_{gs}^2 \right.$$

Figure 4.5 shows the comparison between logarithm of second derivative of

current with respect to gate voltage in model and the measured amplitude of 1st order

harmonic from single tone excitation measurement. The reason we are taking

logarithm of $\left. d^2 I_{ds} \middle/ dV_{gs}^2 \right.$ is because the amplitude of 1st order harmonic is measured in

dBm, which is a logarithm scale unit. Still, in the model, calculation is based on

individual carbon nanotube. In Figure 4.5 (a), which is the nonlinearity calculated

from the model, the important information is that, the value of $log\left(\left. d^2 I_{ds} \middle/ dV_{gs}^2 \right. \right)$ is

large near small absolute value of gate to source voltage, and at higher gate bias, the

linear relationship dominates. Compare this to Figure 4.5 (b), which is the measured

1st order harmonic amplitude. The single tone measurement uses a device having

identical device parameters with the modeling. And the input power is 5dBm at

10GHz. The measured 1st order harmonic has the frequency of 20GHz. The larger

amplitude shows near small absolute value of gate to source voltage and at larger gate

bias, the 1st order harmonic is suppressed to relatively small value, around   -55dBm.
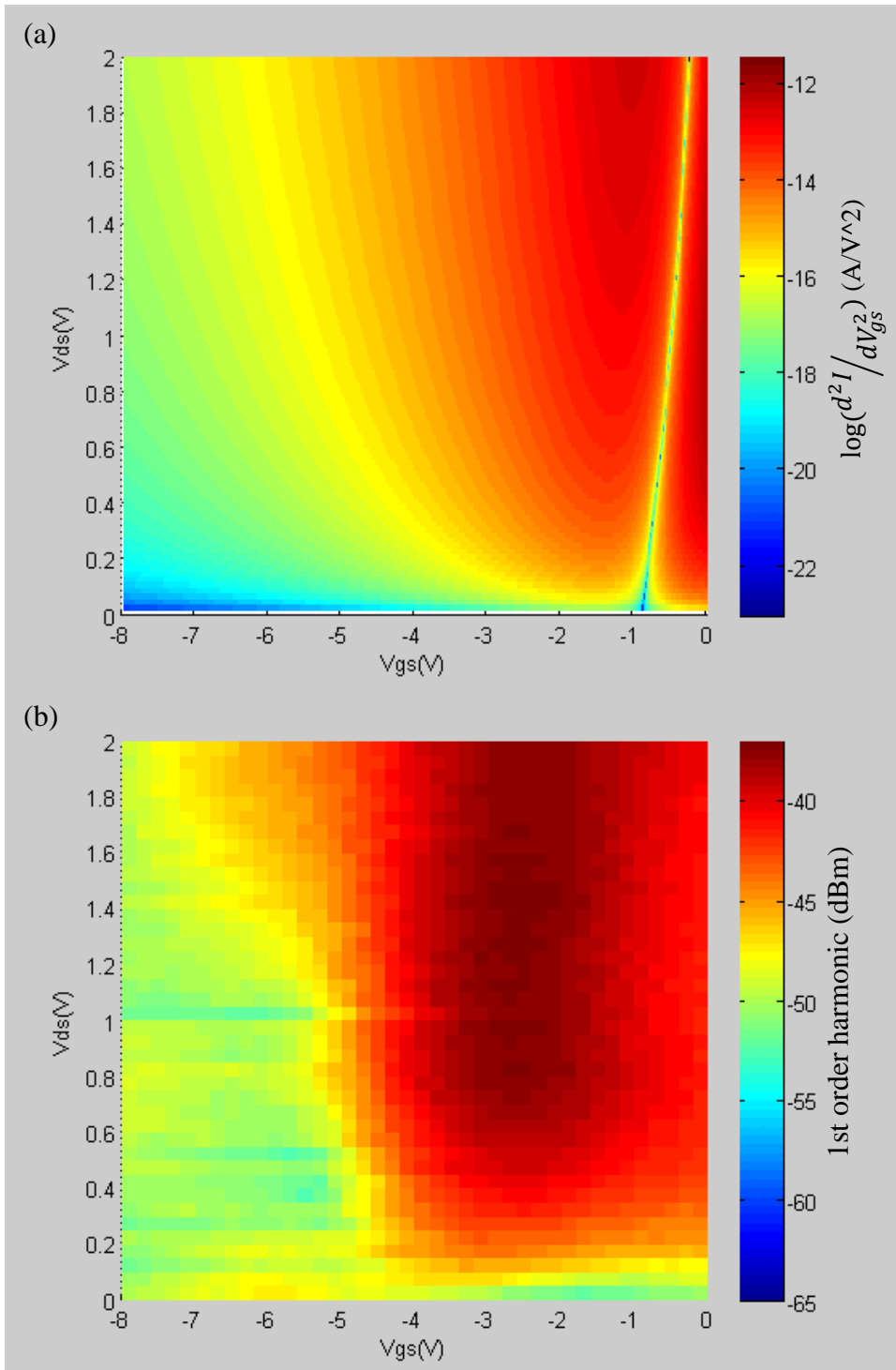
Figure 4.5 Comparison between (a) logarithm of second derivative of current with respect to gate voltage in model $\log(d^2I/dV_{gs}^2)$ and (b) the measured amplitude of 1st

order harmonic.

In terms of the difference between the two figures, the obvious aspect is the scale. Because the amplitude of $d^2 I_{ds} / dV_{gs}^2$ is calculated based on individual nanotube, in order to get the total value for the CNTFET, the number of nanotubes in the channel should be multiplied, which will increase the amplitude. On the other hand, the harmonic measurement measures power, so if we want to convert Figure (a) to (b) with the same unit and scale, we need to multiply v_{gs}^2 in order to get the second term in the small signal equation, and then, multiply v_{gs} again to get the power. We also need to take phase shift when calculating the power. The value of v_{gs} is relatively small, so the multiplication will decrease the amplitude in (a).

Another feature shown in (a) but not in (b) is the small amplitude vertical line started around $V_{gs} = -1V$. There are two possible reasons that it does not appear in (b). The first reason is due to the resolution of the mapping test, we have step of 0.2V in the single tone excitation measurement, and this is much larger than the gate voltage step in the model, which is 0.01V. So the measurement will not be able to show small features. The second reason is the assumption we made about which term in the derivative sequence of transconductance will contribute to the amplitude of 1st order harmonics. Remember we are only taking the 2nd order term in the small signal equation into consideration, which means the contribution to 1st harmonic amplitude from higher order terms are neglected. But if we include the contribution from these higher order terms, it is very likely that they bring in small contribution around the small amplitude line near Figure (a). Since the data is plotted in logarithm scale, a

fraction of contribution will end up bring the blue and green low amplitude parts into

yellow and orange high amplitude parts.

Chapter 5: High frequency mixing based on CNTFETs

5.1. Two-Tone Mixing theory

Two-tone mixing measurement is a powerful measurement setup to examine the nonlinearity of devices can help us extract useful figures of merit for the purpose of RF applications.

Instead of one signal input into the device, two signals with a slight difference in frequency are applied. Name the first frequency as f_1 and the second frequency as f_2, and the difference in frequency between the two signals as δf. First consider the case where the amplitudes of the two signals are the same, and a normal output spectrum of the two tone mixing test will be like in the following figure.

Figure 5.1 shows the output spectrum of a two tone measurement on our CNTFETs. The device under test has a channel length of $3\mu m$, a channel width of $100\mu m$ and a gate length of $2.4\mu m$. The two signals are with the same amplitude of 10dBm and at different frequencies, 0.95GHz and 1.05GHz respectively. Among all the signals in the output spectrum, only the two signals at 0.95GHz and 1.05GHz are the direct output from input signals. All the other signals are either harmonic generated from the fundamental signals or intermodulation signals.

If we follow the same way as the derivation of harmonic generation, where we can get the output small AC signal amplitude as a function of the input signal, the equation will be

$$v_{out} = a_1 v_{in} + a_2 v_{in}^2 + a_3 v_{in}^3 + \cdots$$

where a_n is a function of the partial derivative of current in the channel to the gate voltage. If now we substitute v_{in} with two different signals, say
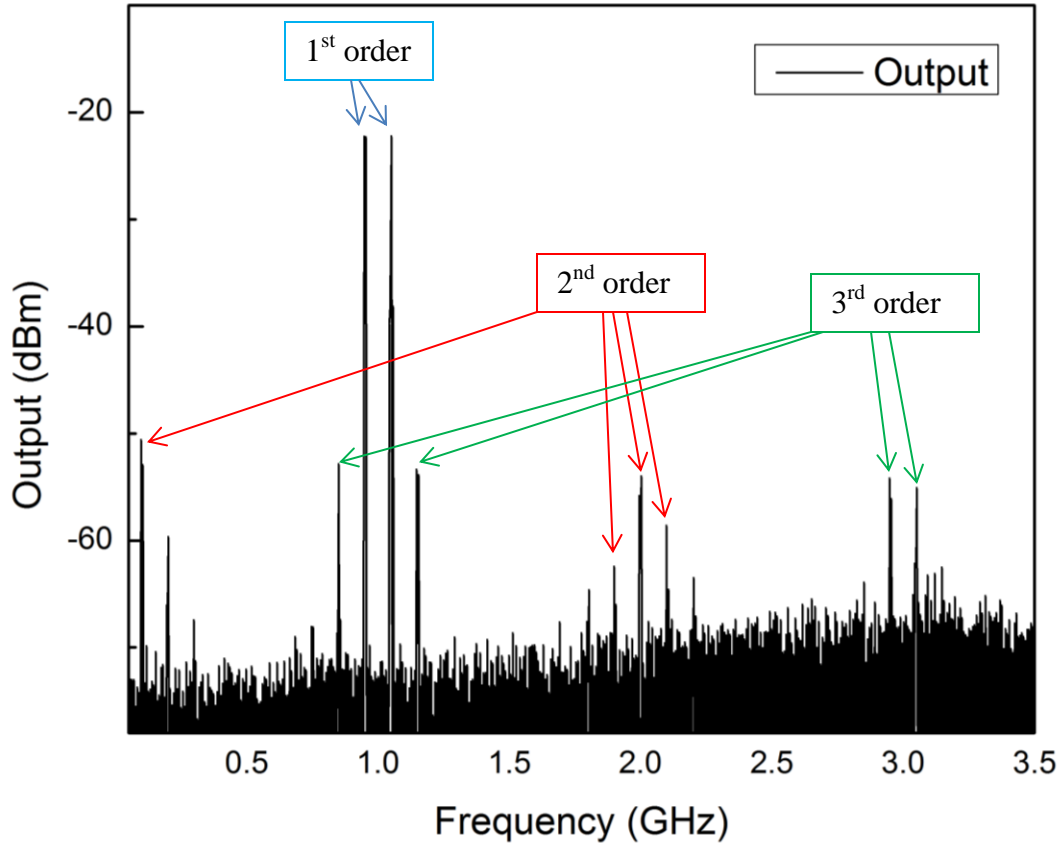
Figure 5.1 The output spectrum of two-tone mixing experiment based on a CNTFET, where the two input frequencies are 0.95GHz and 1.05GHz

$$v_{in} = v_1 \cos(2\pi f_1 t) + v_2 \cos(2\pi f_2 t)$$

and what we get for the expansion of v_{out} will be a sequence of cosine waves.

However, since the a_n sequence tends to converge to zero very quickly, in the

following derivation, we will limit the order of a_n not larger than 3. The expansion

will contain signals with frequencies of $mf_1 + nf_2$, and (m, n) can be any integers,

Table 5.1 shows the amplitude of different sets of frequencies up to the order of 3,

this chart is adapted from Ref[56].

Now if we look back at Figure 5.1, we can label the order of the signals. The

input signals at 0.95GHz and 1.05GHz are f_1 and f_2 respectively. And the second

order terms are the double frequency peaks and the sum and difference of frequency terms. The third order terms are also labeled. One thing to notice is that, the order is different than what we talked about in single tone harmonic generation test. In single tone harmonic excitation test, the 1st order harmonic corresponds to signal with double frequency as the fundamental signal, and the 2nd order harmonic corresponds to signal with triple frequency as the fundamental signal. While in the two-tone mixing test, the order corresponds to the expression of v_{out} as a function of v_{in}. So in Figure 5.1, the 1st order terms represent that they are coming from the linear transformation of input signal, and 2nd order corresponds to the square term of v_{in}. Carry this difference in mind and do not get confused when we talk about 2nd order term in this chapter.

Table 5.1 Output spectrum of a nonlinear amplifier

Order	Frequency	Amplitude	Frequency	Amplitude
1	f_1	$a_1 v_1$	f_2	$a_1 v_2$
2	$2f_1$	$\frac{1}{2} a_2 v_1^2$	$2f_2$	$\frac{1}{2} a_2 v_2^2$
2	$f_1 + f_2$	$a_2 v_1 v_2$	$|f_1 - f_2|$	$a_2 v_1 v_2$
3	$3f_1$	$\frac{1}{4} a_3 v_1^3$	$3f_2$	$\frac{1}{4} a_3 v_2^3$
3	$2f_1 + f_2$	$\frac{3}{4} a_3 v_1^2 v_2$	$|2f_1 - f_2|$	$\frac{3}{4} a_3 v_1^2 v_2$
3	$f_1 + 2f_2$	$\frac{3}{4} a_3 v_1 v_2^2$	$|f_1 - 2f_2|$	$\frac{3}{4} a_3 v_1 v_2^2$

5.2. Two-tone mixing measurement setup

| Signal Generator #1 | Power Combiner | Gate Bias T | CNT FET | S/D Bias T | Spectrum Analyzer |
| Signal Generator #2 | | Source Meter #1 | | | Source Meter #2 |
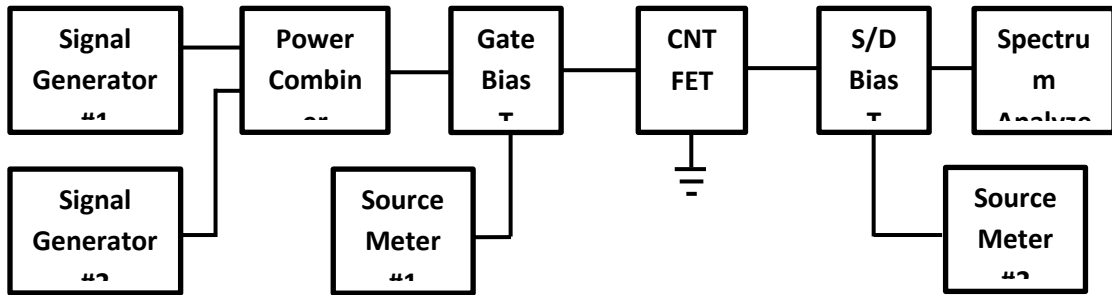
Figure5.2 Schematic diagram of two-tone mixing measurement setup

   Agilent MXG Analog Signal Generator (N5183A) × 2 : 100 kHz ~ 40 GHz

   Agilent EXA Series Spectrum Analyzer (N9010A): 10 Hz ~ 44 GHz

   Keithley 2400 Source-Meters: DC

   Anritsu K250 Bias Tees: DC ~ 40 GHz

   Picoprobe 150 μm pitch microwave probes

The two tone mixing measurement setup we use is shown in Figure 5.3. Similar to

single excitation measurements, the RF signal goes into the gate and the output signal

is read out from the drain, while source is the common ground. The difference from

the single tone excitation measurement is that we now have two signal generators,

and after the two signals are combined via a power combiner, they are combined with

the DC gate voltage. On the drain side, another bias tee is used to change the bias

between source/drain, while the output RF signal is analyzed through a spectrum

analyzer.

        This two tone mixing measurement setup is used in the third order intercept

point measurement and in the characterization of controllable mixing through gate

control with CNTFETs. Compare our setup to a standard mixer, where the two input

signal will act as RF (radio frequency) and LO (local oscillator), while the mixed signal is read from the output spectrum.

For a typical two-tone mixing measurement, two signals with frequency of 4.8GHz and 5.2GHz are generated by the two signal generator, where the 4.8GHz signal is the RF and 5.2GHz signal is the LO. Generally, at least 0dBm amplitude is required in order to observe the harmonics and intermodulation terms in the output spectrum. Two important signals are observed and tracked in the output spectrum, the intermediate frequency (IF) signal which is at 0.4GHz and RL+LO signal, which is at 10GHz. The DC modulation of gate voltage and source/drain bias is V_{gs} from -8V to 8V with a step of 0.2V, and V_{ds} from -2.5V to 2.5V with a step of 0.05V.

5.3. Frequency Mixer

A frequency mixer is a 3-port electronic circuit. Two of the ports are "input" ports and the other port is an "output" port. The ideal mixer "mixes" the two input signals such that the output signal frequency is either the sum (or difference) frequency of the inputs. The nomenclature for the 3 mixer ports are the Local Oscillator (LO) port, the Radio Frequency (RF) port, and the Intermediate Frequency (IF) port. Conceptually, the LO signal acts as the "gate" of the mixer in the sense that the mixer can be considered "on" when the LO is a large voltage and "off" when the LO is a small voltage. The LO port is usually used as an input port.

$$f_{IF} = |f_{LO} - f_{RF}|$$
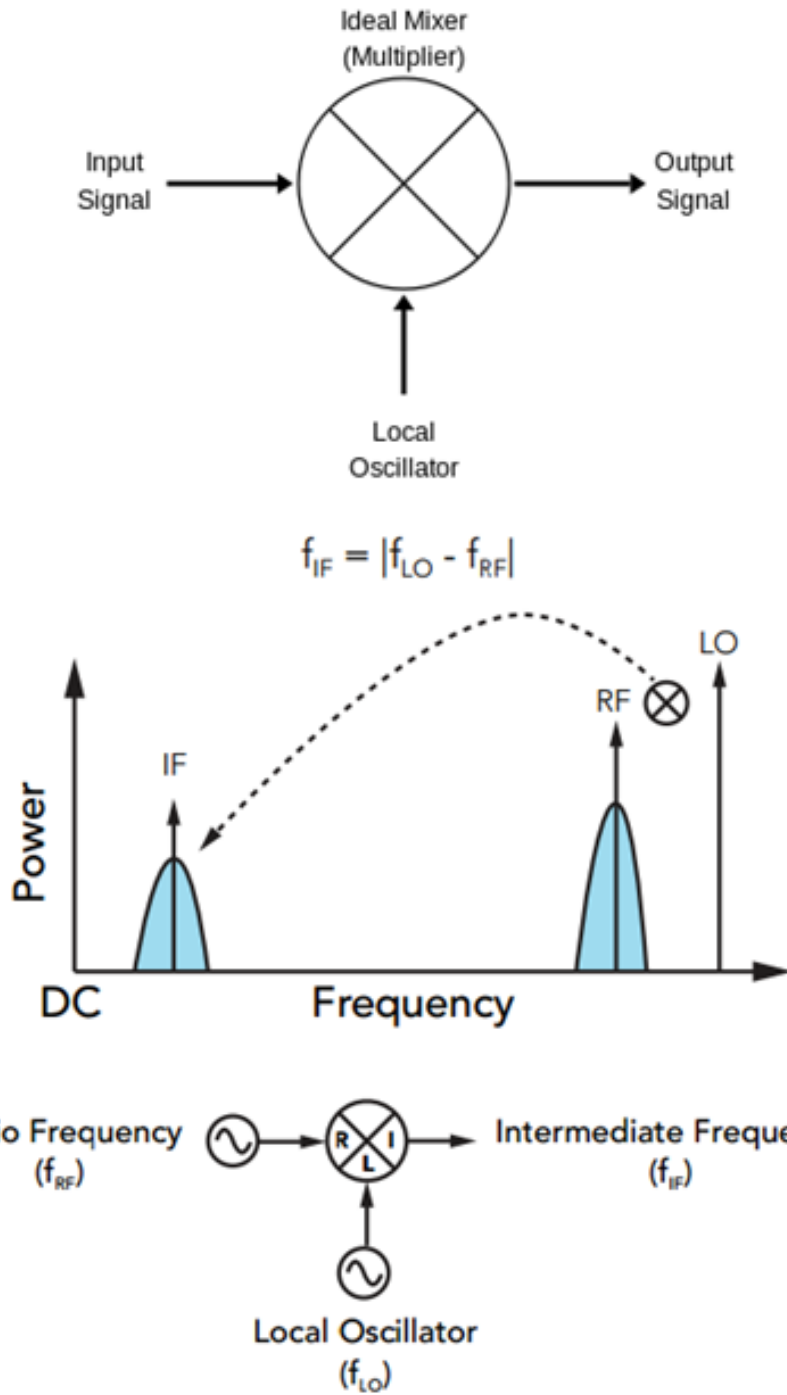
Figure 5.3 (top) inputs and outputs of a frequency mixer (bottom) down-conversion

operation of a frequency mixer

Figure 5.3 shows the inputs and outputs of a frequency mixer on top. And the bottom half shows the down-conversion of a frequency mixer, where the output is the IF signal, and the frequency of the output signal is the difference between LO and RF signal. In principle, any nonlinear device can be used to make a mixer. As it happens, only a few nonlinear devices make "good" mixers. The devices of choice for modern mixer designers are Schottky diodes, GaAs FETs and CMOS transistors. The choice depends on the application. FET and CMOS mixers are typically used in higher volume applications where cost is the main driver.

One important mixer metric is conversion loss. Conversion loss (CL) is defined as the difference in power between the input RF power level and the desired output IF frequency power level. In other words:

$$CL = P_{RF} - P_{IF}$$

where P_{RF} and P_{IF} are in dBm and CL is in dB. For example, if the input RF is -10 dBm and the down-converted IF output signal -17 dBm, then the conversion loss is 7 dB.

In the following measurements, we will treat the higher frequency signal as the LO and the low frequency as RF. And in the output spectrum, we will select the RF+LO signal, which is has the sum of frequency, as the indication of how effective the mixing is.

5.4. Two Tone Mixing Data

Figure 5.4 shows the output spectrum of a two tone mixing measurement results. The device under test is a CNTFET with a channel length of $3\mu m$, a channel width of $100\mu m$ and a gate length of $2.4\mu m$. The two input signals are 0dBm
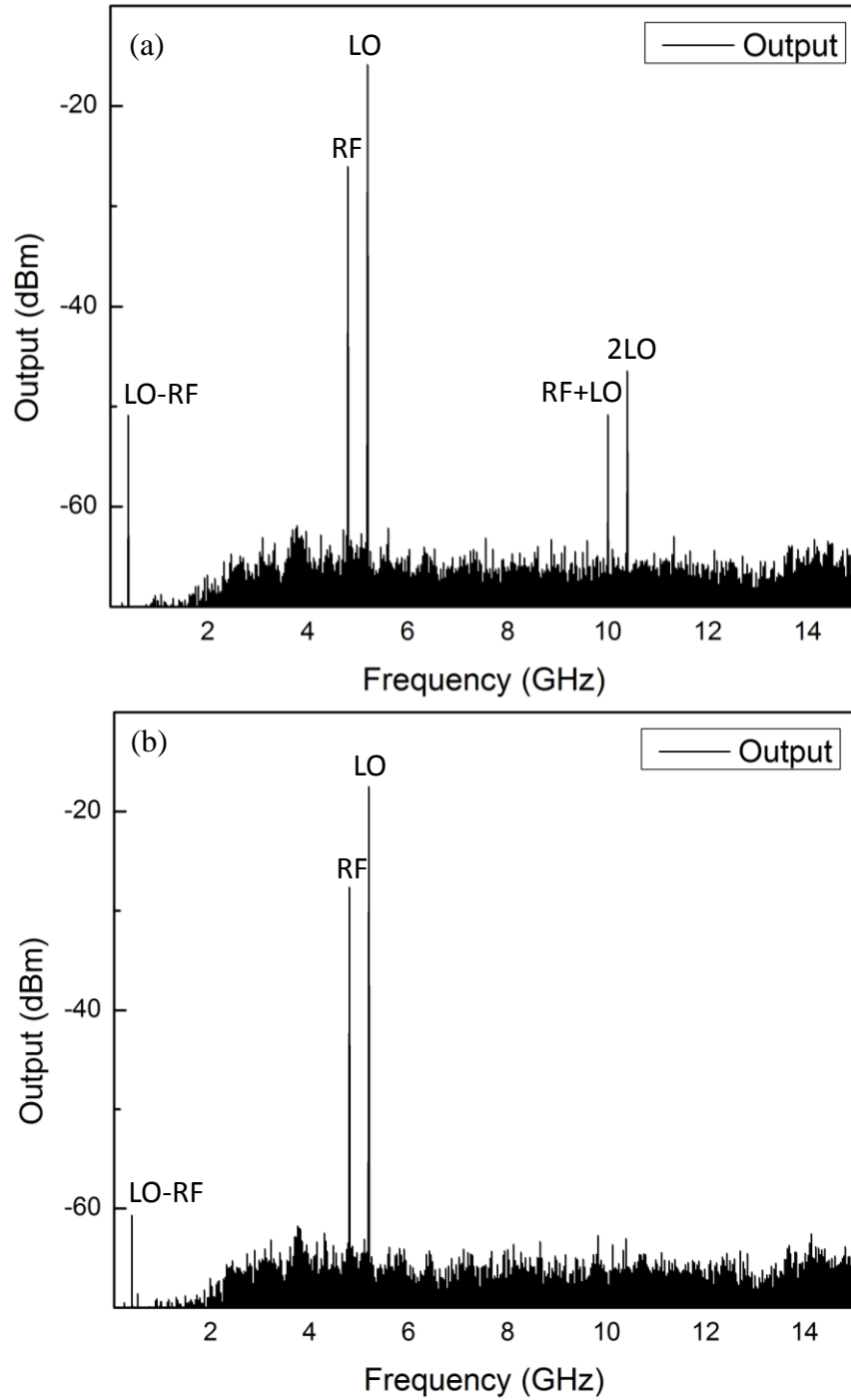
Figure 5.4 Output spectrum of a CNTFET operated as frequency mixer. The DC bias condition is (a) $V_{ds} = 1.5V, V_{gs} = -3V$ and (b) $V_{ds} = 1.5V, V_{gs} = -8V$

amplitude at 4.8GHz and 10dBm at 5.2GHz, and the two signals are labeled as RF

and LO respectively. In the output spectrum, mixed terms besides signals having the

same frequency as the input signals are also labeled.

In Figure 5.4 (a), the DC voltages applied are $V_{ds} = 1.5V, V_{gs} = -3V$. The

output spectrum is a standard output for two-tone mixing with a nonlinear device. The

RF+LO signal is at the frequency of 10GHz and the 2LO signal is at 10.4GHz. The

2RF signal was not obvious in the output spectrum mainly because the amplitude of

RF signal is 0dBm, which is even smaller after passing through the CNTFET. With

larger amplitude of RF signal, the 2RF term will be seen. However, this will not

inflect our observation and comparison since we will take the RF+LO signal as the

main indication of the strength of the mixing.

In Figure 5.4 (b), the DC voltages applied are $V_{ds} = 1.5V, V_{gs} = -8V$.

Compared to (a), the RF+LO and 2LO signals are no longer seen in the output

spectrum, and the LO-RF signal is also 10dBm lower than (a). The reason for the

strong depression of mixed terms is the change of gate voltage. At $V_{gs} = -3V$, the

CNTFET is near the ambipolar region and the nonlinearity of the device is strong,

and as expected, we observe the extra mixed terms in the output spectrum. However,

when we change the gate voltage to $V_{gs} = -8V$, the device is in the p-side linear

region, and the parameters a_n in the output sequence is very small for 2nd or higher

order terms, which leads to the strong suppression of the mixed terms. The following

measurements of RF+LO signal amplitude vs. the gate voltage will further clarify the

relationship.

Figure 5.5 (a) shows the amplitude of RF+LO and 2LO signals vs. the gate

voltage. The device under test is the same as the above measurement, a CNTFET with

a channel length of $3\mu m$, a channel width of $100\mu m$ and a gate length of $2.4\mu m$. The

source/drain bias is set to 1V and gate to source bias is from -10V to 10V. The

important information from the figure is that between the gate voltage range of 6V to

10V and -10V to -5V. The amplitudes of both the signals are very small and down to

the noise floor around -65dBm. At around -5V, the amplitudes of the mixed terms

shows up rapidly, and reach the maximum at around -2V. The difference between the

maximum of mixed signals and the minimum is about 8dBm and 12dBm for RF+LO

and 2LO respectively.

Figure (b) is the transfer curve measured from the same device under the same

bias conditions, and the DC data is taken at the same time with the mixing

measurement. Compared to (a), to the left of the vertical red line, it is the p-type

conduction linear region, and this region corresponds to the negative gate voltage side

mixed signal suppression. Between the red and blue vertical lines, it is the ambipolar

nonlinear conduction region, and the high nonlinearity of the IV relationship

contributes to strong mixing. To the right of the blue vertical line, it is the n-type

conduction linear region, which also suppress mixing. Close to the n-type conduction

region, the transmission between strong mixing and noise floor is not as sharp as it is

on the p-type side, and this can be attribute to the high linearity of the p-type

conduction.
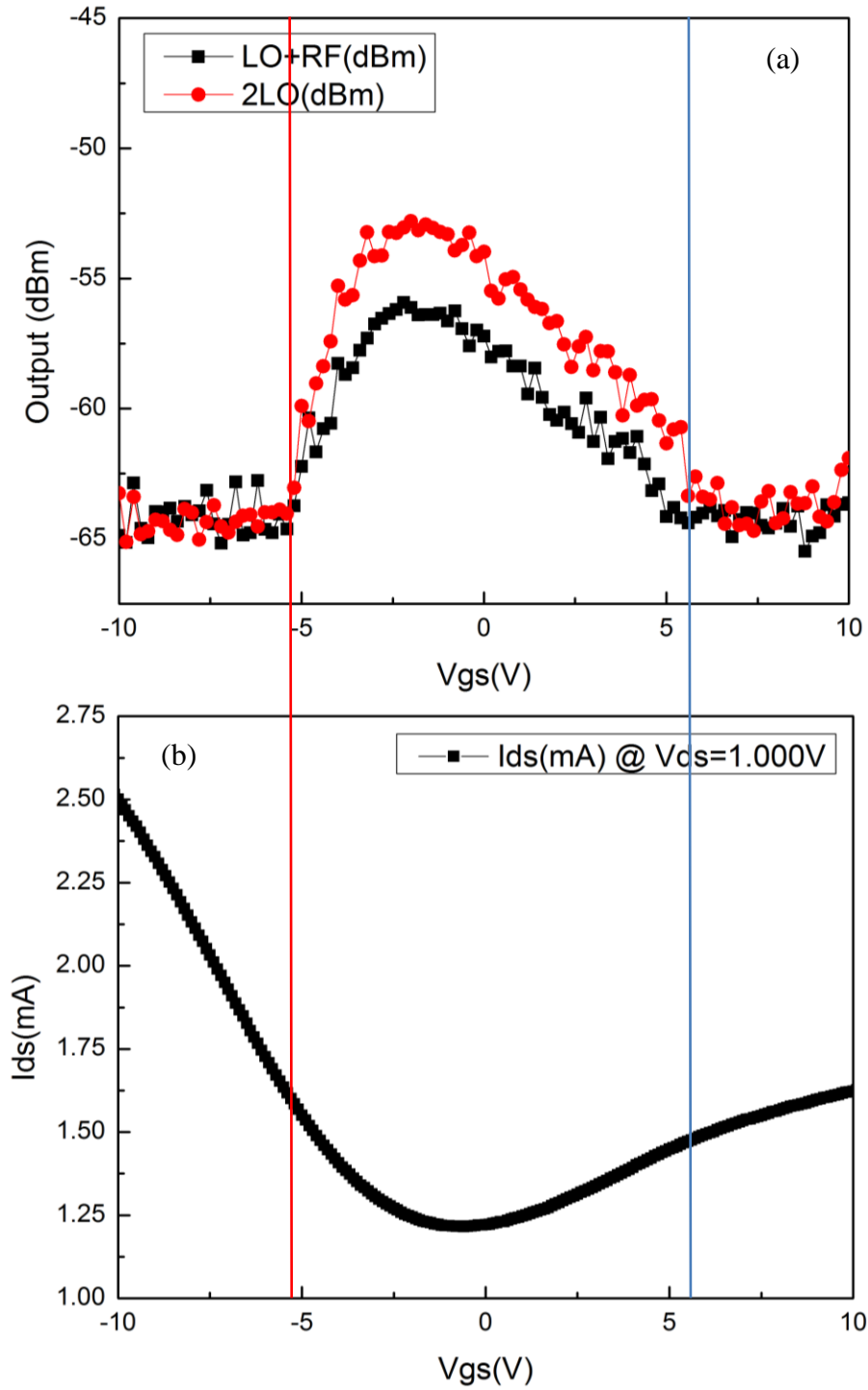
Figure 5.5 (a) LO+RF and 2LO signals amplitude plotted vs. gate voltage (b) transfer curve of the same device under same DC bias conditions

5.5. Controllable radio frequency mixer based on CNTFETs

For a standard mixer, the "gate" of the mixer, which determines whether mixed signals are generated in the output spectrum, is the LO signal amplitude. When the amplitude of LO signal is strong, mixed signals with frequencies at RF+LO and LO-RF can be detected. On the other hand, when the LO signal is turned off, only 2RF frequency signal exists in the output spectrum.

Based on the gate voltage scan measurement, it is reasonable to justify that there are certain regions in the (V_{gs}, V_{ds}) plane where mixing is suppressed, while in other areas, mixing is strong. Figure (a) shows the amplitude of RF+LO mixed signals vs. V_{gs} &V_{ds}. The device is the same CNTFET with a channel length of $3\mu m$, a channel width of $100\mu m$ and a gate length of $2.4\mu m$. Red colors represent large amplitude of RF+LO signal and blue stands for small amplitude. As is expected, the linear regions in transfer curves, which are the far negative and positive gate voltage regions, suppress the mixing. While in the ambipolar regions, mixing is strong. The difference between the strongest mixing signal amplitude and the general background is more than 15dBm. Thus a large difference can be treated as the on and off state of the mixer, and by doing so, we have a method to control the CNTFET mixer via gate voltage instead of using LO signal amplitude. Furthermore, the control from gate voltage over the amplitude of RF+LO signal amplitude means, we are controlling the mixer by changing the DCgate bias instead of changing AC power, and this novel concept of controllable mixing will enable more vesatile radio frequency circuit components and the realization of multifunctional RF device based on CNTs.
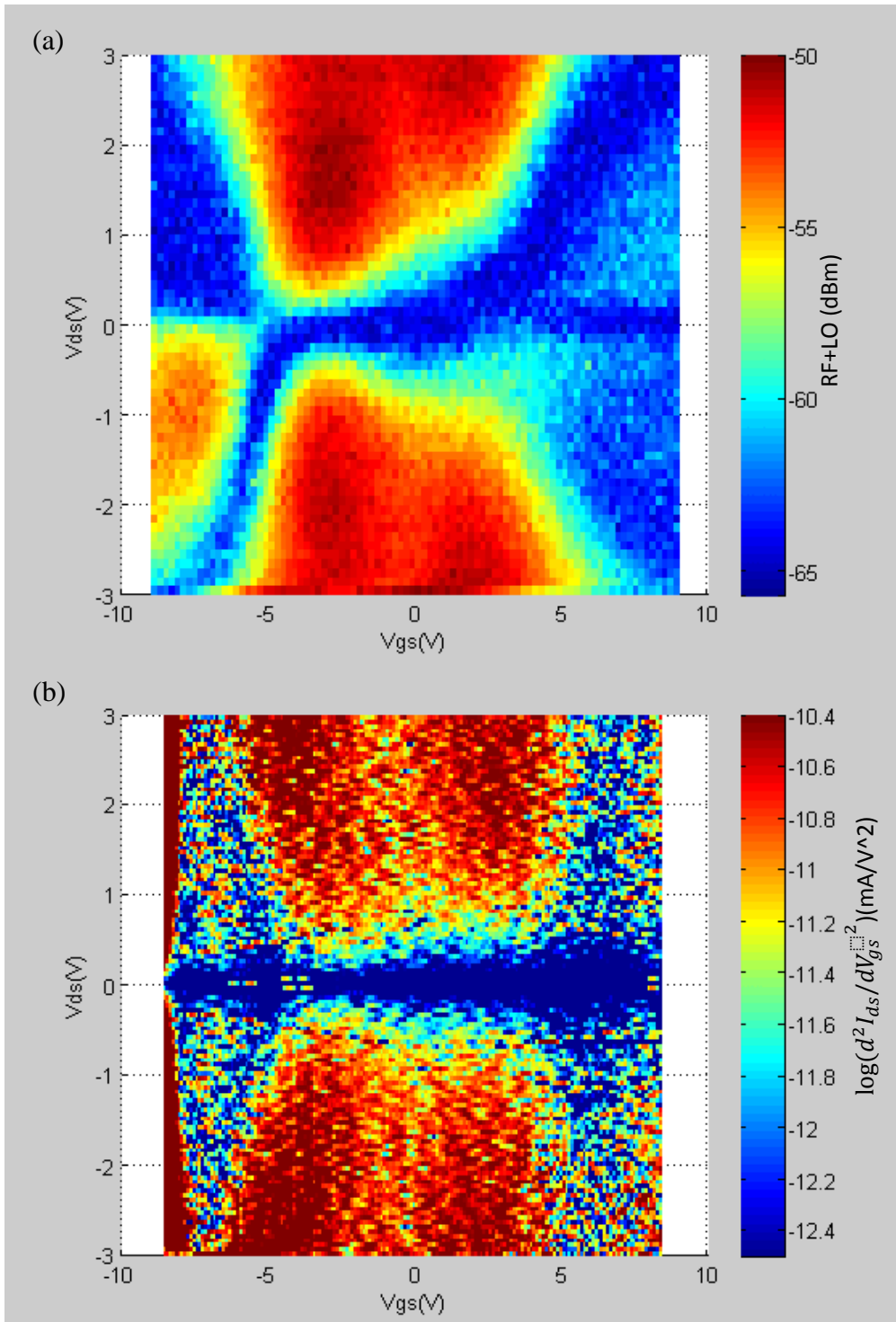
Figure 5.6 (a) Amplitude of RF+LO signal in the output spectrum of two-tone mixing measurement (b) logarithm of the second derivative of current with respect to gate voltage. The comparison between (a) and (b) relates the nonlinearity of transconductance to the generation of intermodulation terms.

An asymmetric region in the negative gate voltage region and negative source/drain bias region is notified. We cannot thoroughly understand the asymmetric behavior. Some of the possible reasons could be the small misplacement of the gate when conducting photolithography or this is due to the device structure. If we want to further understand the cause of the asymmetric behavior, an HFFS model is needed in order to speculate the influence of the device structure on the propagation of electromagnetic waves.

In order to understand the relationship between intermodulation terms and the DC nonlinearity of CNTFETs, we will specifically look at the first three orders in the output spectrum of a frequency mixer. Similar to the harmonic generation theory, we will start with the DC operation equation

$$I_{ds} = f(V_{gs}, V_{ds})$$

Take this equation and do the Taylor expansion around a given voltage operation point with small AC signal (v_{gs}, v_{ds}), and use the definition of transconductance and dynamic conductance,

$$I_{ds} = f(V_{gs} + v_{gs}, V_{ds} + v_{ds})$$

$$= f(V_{gs}, V_{ds}) + g_m \cdot v_{gs} + G_d \cdot v_{ds} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial G_d}{\partial V_{ds}} \cdot \frac{v_{ds}^2}{2!}$$

$$+ g_m G_d \cdot \frac{v_{gs} v_{ds}}{2!} + \cdots$$

Drop all derivative terms related to V_{ds} since we only apply RF signal to gate,

$$I_{ds} = I_{ds0} + g_m \cdot v_{gs} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial^2 g_m}{\partial V_{gs}^2} \cdot \frac{v_{gs}^3}{3!} + \cdots$$

Then we have the small ac signal

$$i_{ds} = g_m \cdot v_{gs} + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{v_{gs}^2}{2!} + \frac{\partial^2 g_m}{\partial V_{gs}^2} \cdot \frac{v_{gs}^3}{3!} + \cdots$$

Now, instead substitute v_{gs} with a single cosine wave, we will put in the sum of two cosine waves with different amplitudes and different frequencies, corresponding to the most general case

$$v_{gs} = A_1 cosw_1 t + A_2 cosw_2 t$$

where A_1 and A_2 are the amplitude of two signals,

$$i_{ds} = g_m \cdot (A_1 cosw_1 t + A_2 cosw_2 t) + \frac{\partial g_m}{\partial V_{gs}} \cdot \frac{(A_1 cosw_1 t + A_2 cosw_2 t)^2}{2!} + \frac{\partial^2 g_m}{\partial V_{gs}^2}$$

$$\cdot \frac{(A_1 cosw_1 t + A_2 cosw_2 t)^3}{3!} + \cdots$$

For the second term

$$\frac{\partial g_m}{\partial V_{gs}} \cdot \frac{(A_1 cosw_1 t + A_2 cosw_2 t)^2}{2!}$$

$$= \frac{1}{2} \frac{\partial g_m}{\partial V_{gs}} \left(A_1{}^2 cos^2 w_1 t + A_2{}^2 cos^2 w_2 t + 2A_1 A_2 cosw_1 t cosw_2 t \right)$$

$$= \frac{1}{2} \frac{\partial g_m}{\partial V_{gs}} \left\{ \frac{1}{2} \left(A_1{}^2 + A_2{}^2 \right) + \frac{1}{2} \left(A_1{}^2 cos2w_1 t + A_2{}^2 cos2w_2 t \right) \right.$$

$$\left. + A_1 A_2 [cos(w_1 - w_2)t + cos(w_1 + w_2)t] \right\}$$

where the $cos(w_1 - w_2)t$ term corresponds to the IF signal of two-tone mixing measurement and $cos(w_1 + w_2)t$ term corresponds to the RF+LO signal from the two-tone mixing. Take out the $cos(w_1 + w_2)t$ term and related coefficients since we are observing RF+LO signal, the absolute value is

$$RF + LO = \frac{1}{2} \frac{\partial g_m}{\partial V_{gs}} A_1 A_2 \, cos(w_1 + w_2)t \quad \propto \quad \frac{d^2 I_{ds}}{dV_{gs}^2}$$

So the amplitude of RF+LO signal is proportional to the second derivative of current with respect to gate/source bias. Actually, not only 2nd order term contains $\cos(w_1 + w_2)t$, all higher even order terms will generate $\cos(w_1 + w_2)t$ term, but the main contribution still comes from the second order term. We will further calculate the third order term to make sure the contribution is from higher order terms but not the third order. The third order term is

$$\frac{\partial^2 g_m}{\partial V_{gs}^2} \cdot \frac{(A_1 cos w_1 t + A_2 cos w_2 t)^3}{3!}$$

$$= \frac{1}{6}\frac{\partial^2 g_m}{\partial V_{gs}^2}\left\{\left(\frac{2}{3}A_1{}^2 A_2 + \frac{3}{4}A_2{}^3\right) cos w_2 t + \left(\frac{2}{3}A_1 A_2{}^2 + \frac{3}{4}A_1{}^3\right) cos w_1 t\right.$$

$$+ \frac{3}{4}A_1 A_2{}^2 \cos(2w_2 - w_1)t + \frac{3}{4}A_1{}^2 A_2 \cos(2w_1 - w_2)t$$

$$+ \frac{3}{4}A_1{}^2 A_2 \cos(2w_1 + w_2)t + \frac{3}{4}A_1 A_2{}^2 \cos(w_1 + 2w_2)t$$

$$\left. + \frac{1}{4}A_1{}^3 cos3w_1 t + \frac{1}{4}A_2{}^3 cos3w_2 t\right\}$$

There is no $\cos(w_1 + w_2)t$ term in the third order expansion, which means the contribution to RF+LO signal besides 2nd order term is from 4th order and higher. And it is reasonable for us to compare the second derivative of current to the amplitude of RF+LO signal.

Based on the calculation above, we further move on to numerically differentiate current with respect to gate voltage and make the comparison between these two. If the similarity is high between these two data sets, it will not only justify the application of generic mixing theory on our CNTFET mixing measurement, but

also deepen our understanding of the working mechanism of our controllable frequency mixer.

Figure 5.6 (b) shows the logarithm of the absolute value of second derivative of I_{ds} with respect to V_{gs}. The method used to extract the second derivative is moving boxcar numerical fitting. The method is illustrated as following. Take the I_{ds} as a whole matrix indexed by V_{gs} and V_{ds}, and for every V_{ds}, I_{ds} is an array indexed by V_{gs}, take the first 5 elements in the array and fit it to a parabola, and take the second order coefficient as the second derivative at the third element point. After that, take second to sixth element and do the same fitting, and the second order coefficient is the second order derivative value for the fourth element. Do the same thing for the whole array of I_{ds} at a certain V_{ds}, and then cover every V_{ds}. This method will not get you the derivatives for the first and last two elements in each array, which means we do not have second order derivatives for the first and last two gate voltage values. The reason we need to use moving boxcar fitting is that the I_{ds} value is from raw data, and taking second order derivative from raw data directly will produce unwanted numerical noise.

Even though moving boxcar method suppresses the noise level in the numerical derivatives, we still get considerable noise in our map of second derivative of I_{ds} with respect to V_{gs}. However, through the noisy data set, we can observe the trend. As expected, the linear transfer region on both p-type and n-type conduction regions have small value of second derivatives. And in the ambipolar region, the second derivative shows up strongly and match the region where the RF+LO signal is larger. Notice that we do not have asymmetric region across source/drain voltage at

negative gate voltage region, which at least confirms that the DC characteristics are not the cause of this asymmetric behavior.

5.6. Measurement of third order intercept point at linear region

There are several ways to confirm that the high linearity of our CNTFETs in the p-type and n-type conduction region. Third order intercept point, also referred to as IP3, is a widely accepted figure of merit to characterize the linearity of a RF transistor. The concept of third order intercept point is that, when the input power increases, both the fundamental signal power and the third order signal power will increase, however, the third order signal power will increase faster. Actually, when scale is in dBm, the third order signal power will increase three times as fast as the fundamental signal power. Since the fundamental signal output power vs. the input power always has a slope of 1, the third order signal power vs. input power will have a slope of 3. And as the input power goes up, compression might happen, that is when the nonlinear transfer characteristics are introduced. However, the linear region extended line of the fundamental signal and the third order signal will intercept with each other at a higher power, and the input power of this point is defined as input power of third order intercept point (IIP3). As is shown in 5.7, fundamental single power intercept with third order signal power at IP3.
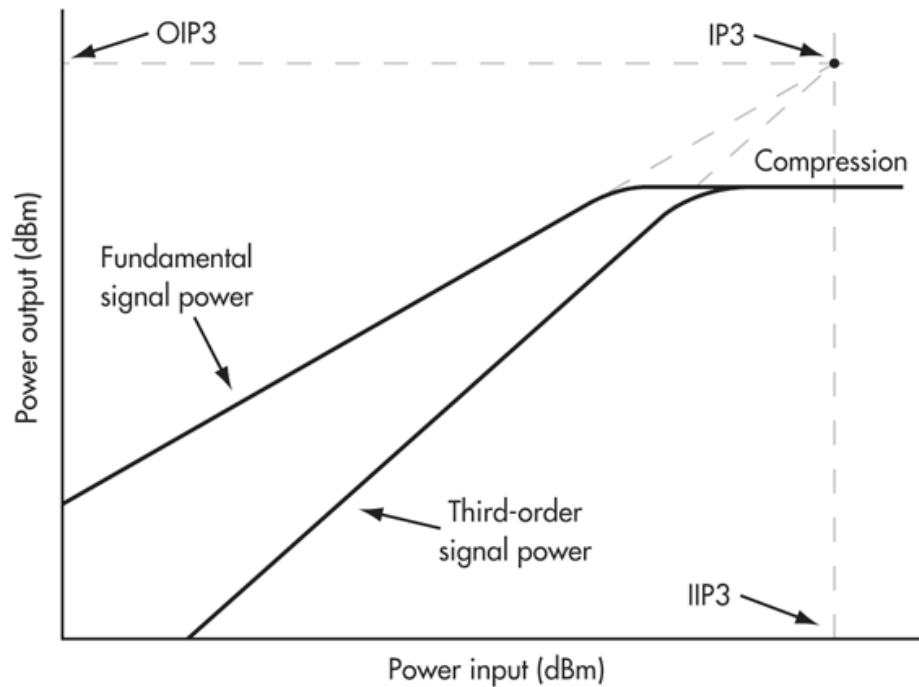
Figure 5.7 Schematic figure showing the concept of IP3, OIP3 and IIP3

The experimental setup of measuring IP3 is the same as the two-tone mixing measurement. The device under test is the same CNTFET with a channel length of $3\mu m$, a channel width of $100\mu m$ and a gate length of $2.4\mu m$. One particular thing about measuring IP3 is that, the two input signals have the same power, with the frequency of f_1 and f_2 ($f_2 > f_1$). The third order output power are measured at the frequency of $2f_2 - f_1$ and $2f_1 - f_2$, as the power of both input signals are increased, the output powers of fundamental signals and third order signals are measured. The two input signals are at 4.8GHz and 5.2GHz, which is identical to that of the two-tone mixing measurement. The two measured third order signals are at 4.6GHz and 5.4GHz. The DC operation point of the measurement is $V_{ds} = 1V, V_{gs} = -6.5V$, which is biased to the linear p-type conduction region in the transfer curve.

The measured fundamental signal and third order signal powers are shown in Figure 5.8. The slope of both the linear fit for the linear parts are close to theoretical values (i.e. 1 and 3). The IIP3 of our device is about 26dBm, which is comparable to CNTFETs with a shorter channel length[20]. This high IP3 confirms that the AC linearity of consistent with our observation on DC transfer curve, and in another way justify the high linearity of our CNTFET when working in the linear transfer region.
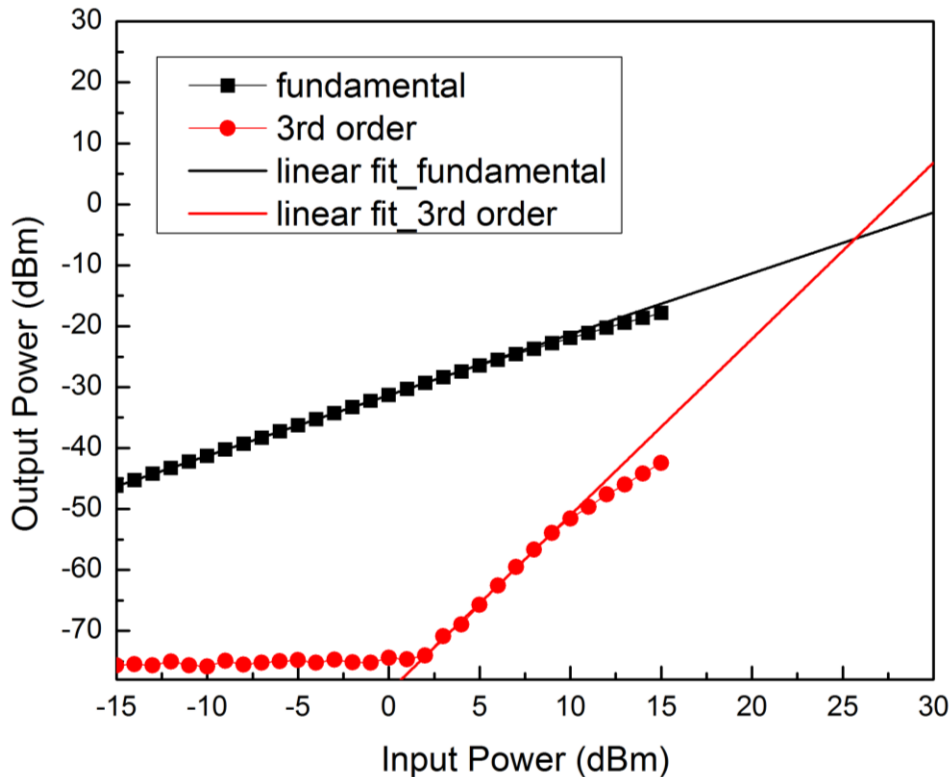
Figure 5.8 Third order interception point measurement data. The black and red dots represent the fundamental and third order signal output power. The black and red lines represent the linear fit of the linear part of fundamental and third order signal output power, respectively.

5.7. Mixing in the 75-110GHz range

Figure 5.9 shows the measurement setup in the 75-110GHz range. It is similar to the

previous mentioned two-tone mixing in the K band range (under 40GHz), the

difference is that the input signals are at higher frequencies, usually between 38GHz

to 40GHz, the harmonics in the output signals are out of the range of our spectrum

analyzer, which has a upper limit of 44GHz. So an external mixer is used to detect

and measure the amplitude of the 1st order harmonics in the 75-100GHz range.

However, the lack of W band bias tees limits our measurement to only passive

mixing, i.e. there is no voltage applied between source and drain.
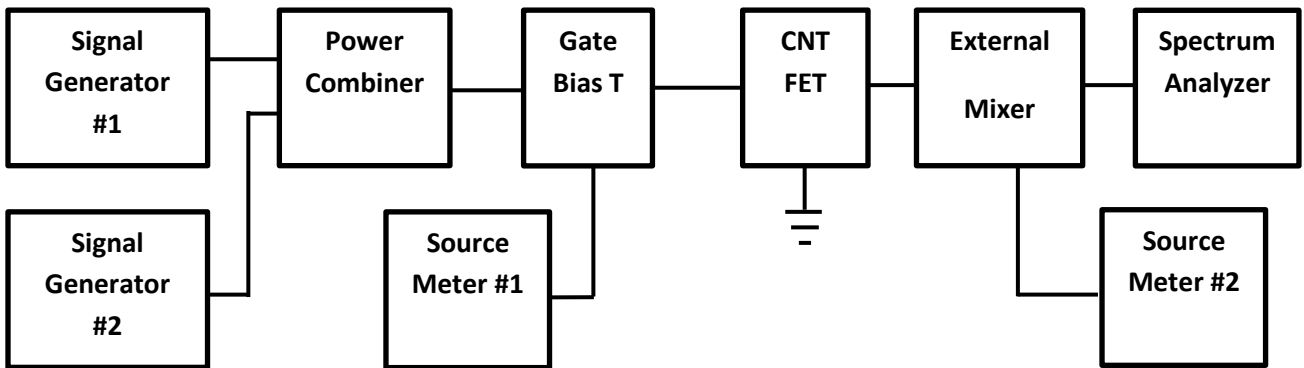
Figure 5.9 Schematic diagram of the experimental setup of mixing measurement in
the 75-110GHz region. Components used are listed below.

Agilent MXG Analog Signal Generator (N5183A) × 2 : 100 kHz ~ 40 GHz

Agilent EXA Series Spectrum Analyzer (N9010A): 10 Hz ~ 44 GHz

Agilent M1970 Series Smart Harmonic Mixers: 75-110GHz

Keithley 2400 Source-Meters: DC

Anritsu K250 Bias Tees: DC ~ 40 GHz

Picoprobe 150 μm pitch microwave probes

The two input signals are chosen as 38.5GHz and 39.5GHz. The upper limit of our signal generator output is 40GHz, so this is basically the highest we can pull out. Take the 38.5GHz input signal as RF and the 39.5GHz signal as LO. The power of RF and LO signal is 15dBm and 18dBm. The reason we set the input power so high is because at such a high frequency, the attenuation through the CNTFET mixer is very large, input signal below 10dBm generally cannot produce enough output single power for the observation. Frequency of 2RF, RF+LO and 2LO output signal is 77GHz, 78GHz and 79GHz respectively. The amplitude of 2RF, RF+LO and 2LO are plotted in Figure 5.10. The important information we get from the plot is even at such a high frequency, both harmonics and intermodulation terms are still affected by the gate bias effectively. The difference between the maximum and minimum in the amplitude of RF+LO signal is close to 15dBm. A closer look at the 3 set of data, they are following the same trend, especially between 2LO and LO+RF. The 2RF signal might be too small to show up in the positive gate voltage range. This makes sense if our measurement is active mixing, because from the two-tone mixing theory part, the second order term is

$$\frac{\partial g_m}{\partial V_{gs}} \cdot \frac{(A_1 cos w_1 t + A_2 cos w_2 t)^2}{2!}$$

$$= \frac{1}{2}\frac{\partial g_m}{\partial V_{gs}}\left(A_1{}^2 cos^2 w_1 t + A_2{}^2 cos^2 w_2 t + 2A_1 A_2 cos w_1 t cos w_2 t\right)$$

$$= \frac{1}{2}\frac{\partial g_m}{\partial V_{gs}}\left\{\frac{1}{2}\left(A_1{}^2 + A_2{}^2\right) + \frac{1}{2}\left(A_1{}^2 cos2w_1 t + A_2{}^2 cos2w_2 t\right)\right.$$

$$\left. + A_1 A_2 [cos(w_1 - w_2)\, t + cos(w_1 + w_2)\, t]\right\}$$

In the equation, $cos2w_1t$, $cos2w_2t$ and $\cos(w_1 + w_2)t$ are sharing the same

coefficient $\frac{\partial g_m}{\partial V_{gs}}$, so that the corresponding 2RF, 2LO and RF+LO will follow the same

trend when tuning gate voltage.
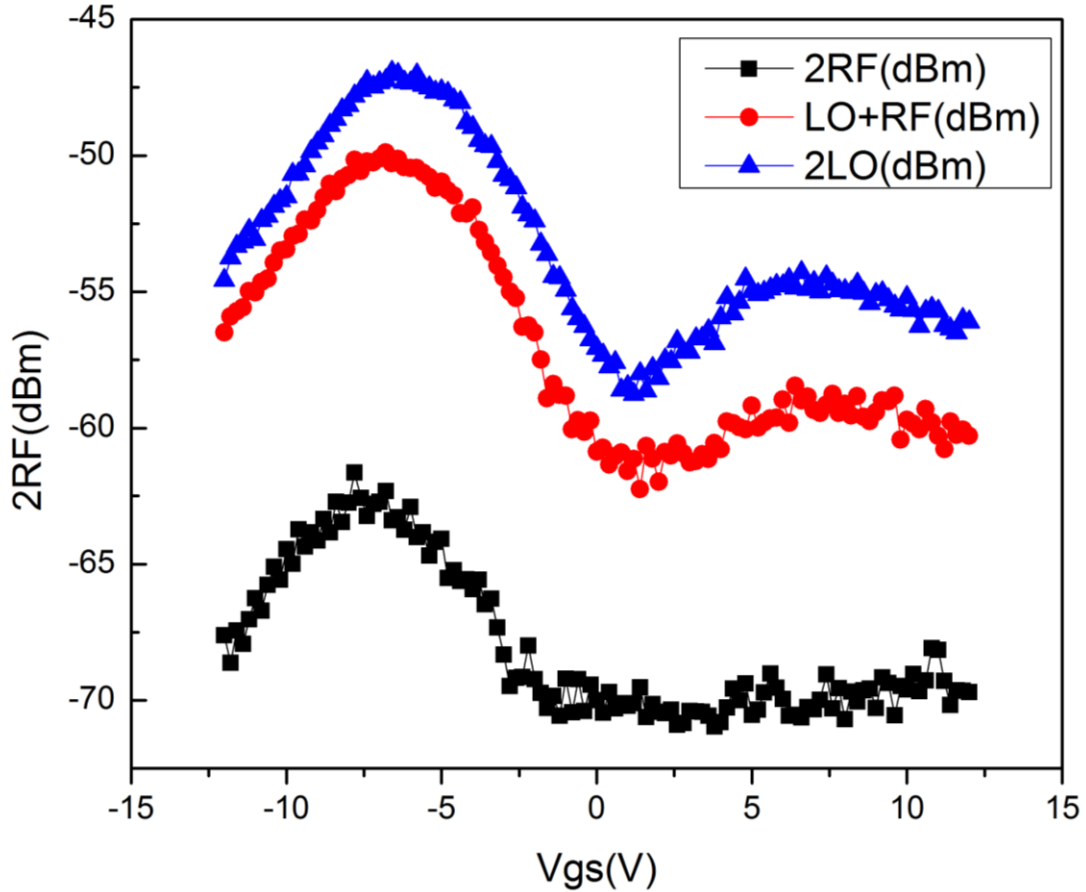
Figure 5.10 Amplitude of harmonics and intermodulation terms of passively two-tone mixing measurement in the 75-110GHz range

However, our mixing is based on passive CNTFET without source/drain bias.

One possible reason for the similarity to active mixing is the DC component caused

by the applied RF signal. Looking at the second order term expansion, the applied RF

and LO signals will induce a DC current in the channel, and the amplitude is

$$\frac{1}{2}\left(A_1{}^2 + A_2{}^2\right)$$

92

And even though the DC current is caused by the AC signal, it will be influenced by gate voltage since the conductance of the channel will be changed when adjusting gate voltage. Thus the CNTFET is actually working with voltage difference between source and drain due to the DC component from AC signal. In general, our CNTFET is used as a passive frequency mixer at W band (75-110GHz) and this is the first observation of mixing based on carbon nanotube devices to date.

Chapter 6: Summary

6.1. Summary of dissertation

In this dissertation, field effect transistors based on horizontally aligned carbon nanotubes are fabricated and built into various radio frequency circuit components including frequency doubler and active/passive mixer. The passive mixer can work up to the 75-110 GHz, which is the highest frequency range to date. The active mixer can be operated as a novel controllable mixer, through which the mixing can be turned on/off by adjusting the gate voltage.

We first describe the CVD growth of horizontally aligned carbon nanotubes on quartz substrate and the advantage of building radio frequency transistors based on these nanotube arrays. The carbon nanotubes array has a density of ~3 tubes/μm and the majority of the nanotubes are aligned perfectly. Since our device fabrication uses only photolithography, the process is highly scalable. 70 transistors are fabricated together on a single chip of quartz substrate and the only limitation of numbers of transistors is the requirement of processing tools and size of the substrate.In depth DC characterization is conducted for the CNTFETs, which shows different conduction region in the transfer curve. The three regions include p-type, ambipolar and n-type conduction region, with p-/n-type conduction region armed with good linearity and ambipolar region with strong nonlinearity.

We then introduce the single tone excitation experiment and how we use the experimental set-up to study the harmonic generation based on our CNTFETs. The way we conduct single tone excitation experiment is introducing RF signal to gate and analyze the output signal from drain. Source is set to be common ground. Before

94

applying source/drain bias and gate voltage to the transistor, we simply apply RF

signal with the amplitude of 10dBm at 10GHz. Even though the device is working

passively, we can clearly observe the generation of 1st order harmonic at 20GHz and

2nd order harmonic at 30GHz from the output spectrum. Next, we apply gate to source

voltage while still not biasing source/drain, we observe the modulation of the

harmonic amplitude as a function of gate voltage. This experiment is conduct on both

CNTFETs and devices with identical structure but without nanotubes in the channel.

Harmonics are not found from the output spectrum of control devices, which confirm

the harmonics are due to active part of the device, i.e. carbon nanotubes in the

channel. The last set of experiment is applying both gate/source voltage and

source/drain bias, making our device work as an active transistor. We plot the 1st

order harmonic amplitude vs. both V_{ds} and V_{gs} in a 2D "heat map". The mapping

shows that the 1st harmonic is strong within ambipolar region and weak in the p-type

and n-type linear conduction region. Based on the harmonic generation theory, the

main contribution to 1st order harmonic is from the 2nd order term in the small signal

equation with the derivative of transcondutance as coefficient. Comparison between

the amplitude of 1st order harmonic and the derivative of transconductance shows

similarity, strong within ambipolar region and weak in linear transfer region, which

supports the conclusion that the harmonic generation from CNTFETs are mainly from

the nonlinearity of transconductance. Utilizing the ambipolar characteristic, we

further develop a frequency doubler working at 5MHz based on CNTFET. By biasing

the gate/source voltage to the minimum current point on transfer curve, where the

local current-voltage relationship can be approached as a parabola, the fundamental

signal is suppressed and the 1st order harmonic shows up strong. The signal amplitude at double frequency is over 15dBm higher than the fundamental signal, which is sufficiently large as for a frequency doubler.

In order to thoroughly understand the nonlinearity of our device, we then start to build an elementary model of electronic transport in CNTFETs. The model combines Drude model with 1D electron transport, and by making average diameter and effective mass assumptions about our carbon nanotube arrays, we are able to calculate the current carried by individual nanotubes in the channel. DC characteristics from the model also show three different conduction regions in the transfer curve. The difference between the model and measurement lies in the existence of metallic nanotubes in the actual device, which is not taken into consideration since they will not affect the nonlinearity of the device due to no gate control. Besides this, the model can describe the DC operation of the device very well. We then start the comparison between the nonlinearity within the model and the amplitude of first order harmonics from single tone excitation measurement. The second derivative of current with respect to gate voltage is calculated based on the model, the strong part lies around small gate bias region, which is the ambipolar part in actual device. The contour profile of large amplitude in 1st order harmonics is similar to the strong part of second derivative of current in the model. The elementary model points out the origin of nonlinearity in CNTFET, which in our case is the combination of diffusive electronic transport and unique 1D electronic transport in carbon nanotubes.

Since the origin of the nonlinearity is clear, we utilize the control of gate over nonlinearity to build radio frequency mixers. Two-tone mixing experiment and generic mixing theory are introduced. Instead of a single signal flow into the gate, we now introduce two signals with close frequencies into the gate. And name these two signals following mixer terminology as RF and LO. Besides the harmonic generation at 2RF and 2LO frequency, a signal at the frequency of RF+LO is observed, which is the intermodulation term. The most interesting part is we can actually tune the amplitude of RF+LO signal by changing the gate bias, instead of the normal control of RF+LO signal amplitude via the amplitude of LO signal. In the linear conduction region, especially the p-type conduction region, where the linearity is superior, both the harmonics and intermodulation terms are suppressed to noise floor, which corresponds to the "off" state of the mixer. In the ambipolar region, where strong nonlinearity shows up, especially the second order term in the small signal Taylor expansion, the RF+LO and harmonic signals shows up strong, which correspond to the "on" state of the mixer. The difference between the large amplitude of intermodulation term and the "off" state noise floor is larger than 15dBm, which is large enough to be operated as a controllable mixer. Our CNTFET device can also work as a passive mixer at very high frequency. We conduct two-tone mixing test in the 75-110GHz range while changing the gate bias. Similar to lower frequency range operation, gate voltage still have control over the amplitude of intermodulation term even when there is no bias applied between source and drain. This is among the highest frequency range mixing experiment based on CNTFETs.

6.2. Possible future work

In terms of further work, we can expect improvement of radio frequency operation of CNTFETs from several aspects. The first aspect is from the carbon nanotube array itself. (1) Increasing the density of carbon nanotubes in the horizontally arrays will improve the radio frequency behavior of the device dramatically. Because a higher density means larger transconductance and less parasitic capacitance per tube, and this will increase the cut-off frequency of the CNTFET greatly. A higher cut-off frequency will enable us to build frequency doubler working at a higher frequency. Several methods are viable for increasing the carbon nanotube density: multi-cycle growth of carbon nanotubes with H_2O as etchant to amorphous carbon and post-growth transfer printing of carbon nanotubes. It has been demonstrate that these methods can greatly increase the density of carbon nanotubes easily to ~10tubes/µm and even higher to ~30 tubes/µm[81]. (2) It will also be very interesting if we can have control over the chirality of carbon nanotubes in the channel, because we will not need to consider the average effect when building the electronic transport model. Certain chirality corresponds to identical diameter and band structure among nanotubes, which makes generalization from individual nanotube to nanotube arrays more convincing. There are some attempts to control the chirality of carbon nanotubes including vapor phase epitaxy and catalyst nanostructure control[82]. However, these attempts suffer from low carbon nanotube density. So there is still a long way to go in terms of optimize the properties of carbon nanotube arrays.

The second aspect is scale down the device. Our CNTFETs have channel length between 2µm and 10µm. Compare the channel length with the mean free path

of carriers (several hundred nanometers, 300nm is often referenced) in carbon nanotubes sort our device into diffusive transport region. If we can shrink the channel length from micron scale into nanometer scale, say 200nm, the electronic transport in carbon nanotubes will transform into ballistic transport. Based on the inherent linearity analysis, ballistic transport is a prerequisite for superior linearity. Such short channel length is beyond the resolution of processing tools using photolithography in our lab, however, we can use e-beam lithography to fabricate short channel device. A viable device fabrication process will include self-aligned source and drain. Basically, after the growth of carbon nanotubes, e-beam pattern 200nm width Al as gate, and self-oxidation of Al will form ultra-thin dielectric layer. After that, using the Al gate as self-aligned mask, deposit thin layer of Pd (10nm) to form source and drain. A concern with this device structure is the large parasitic capacitance between gate/source and gate/drain. But shorter channel length will compensate and still achieve good radio frequency performance.

At last, in terms of measurement, it will be nice to cover the 40-75 GHz range which is left blank in our measurement. By doing this, we can cover the whole frequency range from sub-GHz to 110GHz. We are not able to conduct mixing at 75-110GHz range due to lack of experimental setup component, which is a bias tee at 75-110GHz range. Even though we lack the data, the controllable mixing by changing gate bias is still expected. Another aspect we can do with measurement is building matching network. Even though the device structure is designed to have a 50 Ω impedance, due to the lack of control over the number of carbon nanotubes in the channel, it is always off the desire value. What we can do is measuring the impedance

of the device at certain frequency, which is center frequency we want to conduct mixing experiment, and based on the impedance of the device, we can design the matching network and conduct fabrication using standard lithography. The matching network will be able to enable more power delivered to the device and increase the radio frequency performance of the device.
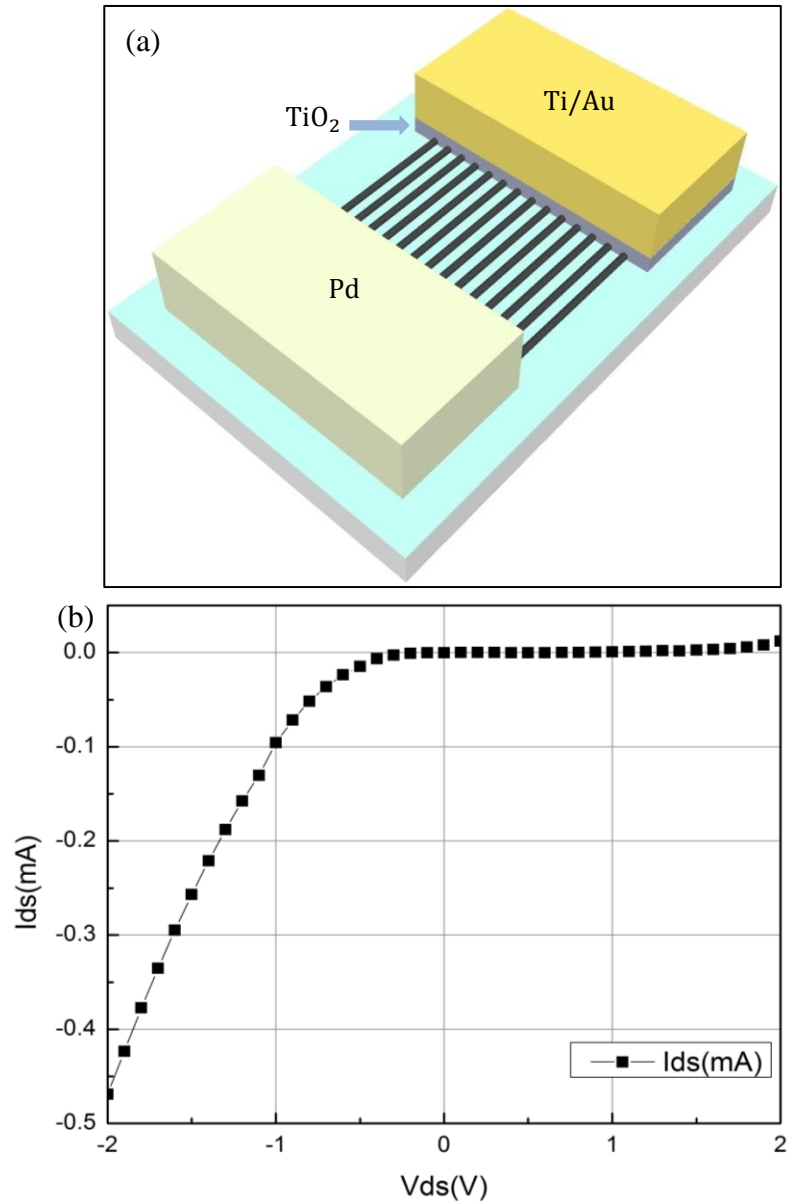
Figure 6.1 (a) Device geometry schematic of a carbon nanotube Schottky diode based on TiO$_2$ hole blocking layer and (b) the current/voltage plot showing rectifying effect

Another interesting experiment we conduct was the fabrication of carbon nanotube

Schottky diode using TiO_2 as hole blocking layer. Figure 6.1 (a) shows the device

geometry schematic of a carbon nanotube Schottky diode based on TiO_2 hole

blocking layer and (b) shows the current/voltage plot. The TiO_2 hole blocking layer

exclude the contribution of metallic nanotubes and making semiconducting nanotubes

sole responsible for carrying the current in the channel[83]. This device geometry not

only allow us to study the current transport proper of semiconducting nanotubes, the

rectifying effect will also allow us to build frequency doubler and mixer working

around $V_{ds} = 0V$. Combined with the gate controllable mixing we demonstrated, we

can build up a carbon nanotube based radio frequency circuit, where we have control

over the mixing product via both V_{ds} and V_{gs}.

Bibliography

[1]   S. Iijima, "Helical microtubules of graphitic carbon," *Nature*, vol. 354, p. 56, 1991.

[2]   X. Wang, Q. Li, J. Xie, Z. Jin, J. Wang, Y. Li, K. Jiang, and S. Fan, "Fabrication of ultralong and electrically uniform single-walled carbon nanotubes on clean substrates," *Nano Lett.*, vol. 9, no. 9, pp. 3137–41, Sep. 2009.

[3]   Q. Wen, R. Zhang, W. Qian, Y. Wang, P. Tan, J. Nie, and F. Wei, "Growing 20 cm Long DWNTs/TWNTs at a Rapid Growth Rate of 80−90 µm/s," *Chem. Mater.*, vol. 22, no. 4, pp. 1294–1296, Feb. 2010.

[4]   P. R. Wallace, "The Band Theory of Graphite," *Phys. Rev.*, vol. 71, no. 9, pp. 622–634, 1947.

[5]   M. Ouyang, J. L. Huang, C. L. Cheung, and C. M. Lieber, "Energy gaps in 'metallic' single-walled carbon nanotubes.," *Science*, vol. 292, no. 5517, pp. 702–5, Apr. 2001.

[6]   T. W. E. and J. M. G. M. M. J. Treacy, "Exceptionally high Young's modulus observed for individual carbon nanotubes," *Nature*, vol. 381, pp. 678–680, 1996.

[7]   S. Berber, Y.-K. Kwon, and D. Tománek, "Unusually High Thermal Conductivity of Carbon Nanotubes," *Phys. Rev. Lett.*, vol. 84, no. 20, pp. 4613–4616, May 2000.

[8]   Z. Yao, C. Kane, and C. Dekker, "High-Field Electrical Transport in Single-Wall Carbon Nanotubes," *Phys. Rev. Lett.*, vol. 84, no. 13, pp. 2941–2944, Mar. 2000.

[9]   Y. Chen and M. S. Fuhrer, "Current-carrying capacity of semiconducting carbon nanotubes," *Phys. Stat. Sol.*, vol. 243, no. 13, pp. 3403–3407, 2006.

[10]   T. Dürkop, S. a. Getty, E. Cobas, and M. S. Fuhrer, "Extraordinary Mobility in Semiconducting Carbon Nanotubes," *Nano Lett.*, vol. 4, no. 1, pp. 35−39, Jan. 2004.

[11]   Q. Cao and J. A. Rogers, "Ultrathin Films of Single-Walled Carbon Nanotubes for Electronics and Sensors : A Review of Fundamental and Applied Aspects," *Adv. Mater.*, vol. 21, pp. 29–53, 2009.

[12]   M. Steiner, M. Engel, Y. Lin, Y. Wu, K. Jenkins, A. A. Green, M. C. Hersam, R. Krupke, and P. Avouris, "High-frequency performance of scaled carbon nanotube array field-effect transistors," *Appl. Phys. Lett.*, no. 101, p. 053123, 2012.

[13]   C. Rutherglen, D. Jain, and P. Burke, "rf resistance and inductance of massively parallel single walled carbon nanotubes : Direct , broadband measurements and near perfect 50 Ω impedance matching," *Appl. Phys. A*, vol. 93, p. 083119, 2008.

[14]   C. Rutherglen, D. Jain, and P. Burke, "Nanotube electronics for radiofrequency applications," *Nat. Nanotechnol.*, vol. 4, no. 12, pp. 811–819, 2009.

[15]   C. Wang, A. Badmaev, A. Jooyaie, M. Bao, K. L. Wang, and W. E. T. Al, "Radio Frequency and Linearity Performance of Transistors Using High-Purity Semiconducting Carbon Nanotubes," *ACS Nano*, vol. 5, pp. 4169–4176, 2011.

[16]   L. Nougaret, H. Happy, G. Dambrine, V. Derycke, J. Bourgoin, and A. A. Green, "80 GHz field-effect transistors produced using high purity semiconducting single-walled carbon nanotubes," *Appl. Phys. Lett.*, vol. 94, p. 243505, 2009.

[17]   C. Kocabas, S. Dunham, Q. Cao, K. Cimino, X. Ho, H. Kim, D. Dawson, J. Payne, M. Stuenkel, H. Zhang, T. Banks, M. Feng, S. V Rotkin, and J. A. Rogers, "High-Frequency Performance of Submicrometer Transistors That Use Aligned Arrays of Single-Walled Carbon Nanotubes," *Nano Lett.*, vol. 9, no. 5, pp. 1937–1943, 2009.

[18]   Z. Wang, S. Liang, Z. Zhang, H. Liu, H. Zhong, L. Ye, S. Wang, W. Zhou, J. Liu, Y. Chen, J. Zhang, and L. Peng, "Scalable Fabrication of Ambipolar Transistors and Radio- Frequency Circuits Using Aligned Carbon Nanotube Arrays," *Adv. Mater.*, vol. 26, pp. 645–652, 2014.

[19]   and L. P. L. Ding, Z. Wang, T. Pei, Z. Zhang, S. Wang, H. Xu, F. Peng, Y. Li, "Self-Aligned U-Gate Carbon Nanotube Field-Effect Transistor with Extremely Small Parasitic Capacitance and Drain-Induced Barrier Lowering," *ACS Nano*, vol. 5, no. 4, pp. 2512–2519, 2011.

[20]   Y. Che, Y. Lin, P. Kim, C. Zhou, and C. H. E. E. T. Al, "T-Gate Aligned Nanotube Radio Frequency Transistors and Circuits with Superior Performance," *ACS Nano*, vol. 7, no. 5, pp. 4343–4350, 2013.

[21]   S. O. Koswatta, A. Valdes-garcia, M. B. Steiner, Y. Lin, S. Member, and P. Avouris, "Ultimate RF Performance Potential of Carbon Electronics," vol. 59, no. 10, pp. 2739–2750, 2011.

[22] S. Hasan, S. Salahuddin, M. Vaidyanathan, and M.A.Alam, "High-Frequency Performance Projections for Ballistic Carbon-Nanotube Transistors," *Ieee Trans. Nanotechnol.*, vol. 5, pp. 14–22, 2006.

[23] L. C. Castro and D. L. Pulfrey, "Extrapolated f max for carbon nanotube field-effect transistors," *Nanotechnology*, vol. 17, pp. 300–304, 2006.

[24] L. C. Castro, D. L. John, S. Member, D. L. Pulfrey, M. Pourfath, A. Gehring, H. Kosina, and A. E. Circuit, "Method for Predicting f T for Carbon Nanotube FETs," *Ieee Trans. Nanotechnol.*, vol. 4, pp. 699–704, 2005.

[25] S. Rosenblatt, H. Lin, V. Sazonova, S. Tiwari, and P. L. McEuen, "Mixing at 50 GHz using a single-walled carbon nanotube transistor," *Appl. Phys. Lett.*, vol. 87, no. 15, p. 153111, 2005.

[26] A. a. Pesetski, J. E. Baumgardner, E. Folk, J. X. Przybysz, J. D. Adam, and H. Zhang, "Carbon nanotube field-effect transistor operation at microwave frequencies," *Appl. Phys. Lett.*, vol. 88, no. 11, p. 113103, 2006.

[27] E. D. Cobas, S. M. Anlage, and M. S. Fuhrer, "Single Carbon Nanotube Schottky Diode Microwave Rectifiers," *IEEE Trans. Electron Devices*, vol. 59, no. 10, pp. 2726–2732, 2011.

[28] S. J.Kang, C. Kocabas, T. Ozel, M. Shim, N. Pimparkar, M. A. Alam, S. V Rotkin, and J. A. Rogers, "High-performance electronics using dense , perfectly aligned arrays of single-walled carbon nanotubes," *Nat. Nanotechnol.*, vol. 2, pp. 230–236, 2007.

[29] S. Chiashi, H. Okabe, T. Inoue, J. Shiomi, T. Sato, S. Kono, M. Terasawa, and S. Maruyama, "Growth of Horizontally Aligned Single-Walled Carbon Nanotubes on the Singular R-Plane (10 − 11) of Quartz," *J. Phys. Chem. C*, vol. 116, pp. 6805–6808, 2012.

[30] D. Yuan, L. Ding, H. Chu, Y. Feng, T. P. McNicholas, and J. Liu, "Horizontally aligned single-walled carbon nanotube on quartz from a large variety of metal catalysts.," *Nano Lett.*, vol. 8, no. 8, pp. 2576–9, Aug. 2008.

[31] L. Huang, B. White, M. Y. Sfeir, M. Huang, H. X. Huang, S. Wind, J. Hone, and S. O'Brien, "Cobalt ultrathin film catalyzed ethanol chemical vapor deposition of single-walled carbon nanotubes.," *J. Phys. Chem. B*, vol. 110, no. 23, pp. 11103–9, Jun. 2006.

[32] F. Gao, L. Zhang, Y. Yang, and S. Huang, "Quality of horizontally aligned single-walled carbon nanotubes: Is methane as carbon source better than ethanol?," *Appl. Surf. Sci.*, vol. 256, no. 11, pp. 3357–3360, Mar. 2010.

[33]   L. Ding, A. Tselev, J. Wang, D. Yuan, H. Chu, T. P. McNicholas, Y. Li, and J. Liu, "Selective growth of well-aligned semiconducting single-walled carbon nanotubes.," *Nano Lett.*, vol. 9, no. 2, pp. 800–5, Feb. 2009.

[34]   N. Patil, A. Lin, S. Member, E. R. Myers, K. Ryu, A. Badmaev, C. Zhou, H. P. Wong, S. Mitra, and S. Member, "Wafer-Scale Growth and Transfer of Aligned Single-Walled Carbon Nanotubes," *IEEE Trans. Electron Devices*, vol. 8, no. 4, pp. 498–504, 2009.

[35]   I. Ibrahim, A. Bachmatiuk, J. H. Warner, B. Büchner, G. Cuniberti, and M. H. Rümmeli, "CVD-Grown Horizontally Aligned Single-Walled Carbon Nanotubes : Synthesis Routes and Growth Mechanisms," *small*, vol. 8, pp. 1973–1992, 2012.

[36]   M. A. Topinka, M. W. Rowell, D. Goldhaber-gordon, M. D. Mcgehee, D. S. Hecht, and G. Gruner, "Charge Transport in Interpenetrating Networks of Semiconducting and Metallic Carbon Nanotubes," *Nano Lett.*, vol. 9, pp. 1866–1871, 2009.

[37]   V. A. Online, S. Park, and Z. Bao, "A review of fabrication and application of carbon nanotube film-based electronics," *Nanoscale*, vol. 5, pp. 1727–1752, 2013.

[38]   G. Zhang, P. Qi, X. Wang, Y. Lu, X. Li, R. Tu, S. Bangsaruntip, D. Mann, L. Zhang, and H. Dai, "Selective etching of metallic carbon nanotubes by gas-phase reaction.," *Science*, vol. 314, no. 5801, pp. 974–7, Nov. 2006.

[39]   P. G. Collins, M. S. Arnold, and P. Avouris, "Engineering Carbon Nanotubes and Nanotube Circuits Using Electrical Breakdown," *Science (80-.).*, vol. 292, pp. 706–709, 2001.

[40]   S. Shekhar, M. Erementchouk, M. N. Leuenberger, and S. I. Khondaker, "Correlated electrical breakdown in arrays of high density aligned carbon nanotubes," *Appl. Phys. Lett.*, vol. 98, p. 243121, 2011.

[41]   P. Li and J. Zhang, "Sorting out semiconducting single-walled carbon nanotube arrays by preferential destruction of metallic tubes using water," *J. Mater. Chem.*, vol. 21, no. 32, p. 11815, 2011.

[42]   B. K. Sarker, S. Shekhar, and S. I. Khondaker, "Semiconducting Enriched Carbon Nanotube Aligned Arrays of Tunable Density and Their Electrical Transport Properties," *ACS Nano*, vol. 5, pp. 6297–6305, 2011.

[43]   S. H. Jin, S. N. Dunham, J. Song, X. Xie, J. Kim, C. Lu, A. Islam, F. Du, J. Kim, J. Felts, Y. Li, F. Xiong, M. A. Wahab, M. Menon, E. Cho, K. L. Grosse, D. J. Lee, H. U. Chung, E. Pop, M. A. Alam, W. P. King, Y. Huang, and J. A.

Rogers, "Using nanoscale thermocapillary flows to create arrays of purely semiconducting single-walled carbon nanotubes," *Nat. Nanotechnol.*, vol. 8, pp. 347–355, 2013.

[44]   A. Tunnell, V. Ballarotto, and J. Cumings, "The selective removal of metallic carbon nanotubes from As-grown arrays on insulating substrates," *Appl. Phys. Lett.*, vol. 101, p. 198109, 2012.

[45]   J. Kim, B. Lee, S. Lee, and G. Jeong, "Size engineering of metal nanoparticles to diameter-specified growth of single-walled carbon nanotubes with horizontal alignment on quartz," *Nanotechnology*, vol. 23, p. 105607, 2012.

[46]   G. Hong, Y. Chen, P. Li, and J. Zhang, "Controlling the growth of single-walled carbon nanotubes on surfaces using metal and non-metal catalysts," *Carbon N. Y.*, vol. 50, no. 6, pp. 2067–2082, May 2012.

[47]   W. Zhou, S. Zhan, L. Ding, and J. Liu, "General Rules for Selective Growth of Enriched Semiconducting Single Walled Carbon Nanotubes with Water Vapor as in Situ Etchant," *J. Am. Chem. Soc.*, vol. 134, pp. 14019–14026, 2012.

[48]   D. Shi, W. D. Tennyson, J. C. Keay, E. S. Sanchez, M. B. Johnson, and D. E. Resasco, "Role of water on the surface-guided growth of horizontally aligned single-walled carbon nanotubes on quartz," *Chem. Phys. Lett.*, vol. 525–526, pp. 82–86, Feb. 2012.

[49]   T. Yamada, A. Maigne, M. Yudasaka, K. Mizuno, D. N. Futaba, M. Yumura, S. Iijima, and K. Hata, "Revealing the Secret of Water-Assisted Carbon Nanotube Synthesis by Microscopic Observation of the Interaction of Water on the Catalysts," *Nano Lett.*, vol. 8, pp. 4288–4292, 2008.

[50]   J. Appenzeller, J. Knoch, V. Derycke, R. Martel, S. Wind, and P. Avouris, "Field-Modulated Carrier Transport in Carbon Nanotube Transistors," *Phys. Rev. Lett.*, vol. 89, no. 12, p. 126801, Aug. 2002.

[51]   M. Burghard, H. Klauk, and K. Kern, "Carbon-Based Field-Effect Transistors for Nanoelectronics," *Adv. Mater.*, vol. 21, pp. 2586–2600, 2009.

[52]   C. Kocabas, N. Pimparkar, O. Yesilyurt, S. J. Kang, M. a Alam, and J. a Rogers, "Experimental and theoretical studies of transport through large scale, partially aligned arrays of single-walled carbon nanotubes in thin film type transistors.," *Nano Lett.*, vol. 7, no. 5, pp. 1195–202, May 2007.

[53]   X. Wang, Y. Ouyang, X. Li, H. Wang, J. Guo, and H. Dai, "Room-Temperature All-Semiconducting Sub-10-nm Graphene Nanoribbon Field-Effect Transistors," *Phys. Rev. Lett.*, vol. 100, no. 20, p. 206803, May 2008.

[54] Z. Yin, H. Li, H. Li, L. Jiang, Y. Shi, Y. Sun, G. Lu, Q. Zhang, X. Chen, and H. Zhang, "Single-Layer MoS2 Phototransistors," *ACS Nano*, vol. 6, pp. 74–80, 2012.

[55] A. Javey, J. Guo, Q. Wang, M. Lundstrom, and H. Dai, "Ballistic carbon nanotube field-effect transistors," *Science (80-.).*, vol. 424, pp. 654–657, 2003.

[56] J. E. Baumgardner, A. A. Pesetski, J. M. Murduck, J. X. Przybysz, J. D. Adam, and H. Zhang, "Inherent linearity in carbon nanotube field-effect transistors Inherent linearity in carbon nanotube field-effect transistors," *Appl. Phys. Lett.*, vol. 91, p. 052107, 2007.

[57] M. S. Purewal, B. H. Hong, A. Ravi, B. Chandra, J. Hone, and P. Kim, "Scaling of Resistance and Electron Mean Free Path of Single-Walled Carbon Nanotubes," *Phys. Rev. Lett.*, vol. 98, p. 186808, 2007.

[58] S. Ilani, L. A. K. Donev, M. Kindermann, and P. L. Mceuen, "Measurement of the quantum capacitance of interacting electrons in carbon nanotubes," *Nat. Phys.*, vol. 2, pp. 687–691, 2006.

[59] M. Zhang, L. L. Brooks, N. Chartuprayoon, W. Bosze, Y. Choa, and N. V Myung, "Palladium/single-walled carbon nanotube back-to-back Schottky contact-based hydrogen sensors and their sensing mechanism.," *ACS Appl. Mater. Interfaces*, vol. 6, no. 1, pp. 319–26, Jan. 2014.

[60] J. Svensson and E. E. B. Campbell, "Schottky barriers in carbon nanotube-metal contacts," *Appl. Phys. Lett.*, vol. 110, p. 111101, 2011.

[61] X. Ho, L. Ye, S. V Rotkin, X. Xie, F. Du, S. Dunham, and J. Zaumseil, "Theoretical and Experimental Studies of Schottky Diodes that Use Aligned Arrays of Single-Walled Carbon Nanotubes," *Nano Res.*, vol. 3, pp. 444–451, 2010.

[62] V. Perebeinos, J. Tersoff, and W. Haensch, "Schottky-to-Ohmic Crossover in Carbon Nanotube Transistor Contacts," *Phys. Rev. Lett.*, vol. 111, no. 23, p. 236802, Dec. 2013.

[63] M. a. Hughes, K. P. Homewood, R. J. Curry, Y. Ohno, and T. Mizutani, "An ultra-low leakage current single carbon nanotube diode with split-gate and asymmetric contact geometry," *Appl. Phys. Lett.*, vol. 103, no. 13, p. 133508, 2013.

[64] X. Ho, L. Ye, S. V Rotkin, Q. Cao, S. Unarunotai, S. Salamat, M. A. Alam, and J. A. Rogers, "Scaling Properties in Transistors That Use Aligned Arrays of Single-Walled Carbon Nanotubes," *Nano Lett.*, vol. 10, pp. 499–503, 2010.

[65] Z. Wang, Z. Zhang, H. Zhong, T. Pei, S. Liang, L. Yang, S. Wang, and L. Peng, "Carbon Nanotube Based Multifunctional Ambipolar Transistors for AC Applications," *Adv. Funct. Mater.*, vol. 23, pp. 446–450, 2013.

[66] Z. Wang, L. Ding, T. Pei, Z. Zhang, S. Wang, T. Yu, X. Ye, F. Peng, Y. Li, and L. Peng, "Large Signal Operation of Small Band-Gap Transistor : A High-Performance Frequency Doubler," *Nano Lett.*, vol. 10, pp. 3648–3655, 2010.

[67] S. G. Shirazi and S. Mirzakuchaki, "High on / off current ratio in ballistic CNTFETs based on tuning the gate insulator parameters for different ambient temperatures," *Appl. Phys. A*, vol. 113, pp. 447–457, 2013.

[68] Z. Zhong, N. M. Gabor, J. A. Y. E. Sharping, A. L. Gaeta, and P. L. Mceuen, "Terahertz time-domain measurement of ballistic electron resonance in a single-walled carbon nanotube," *Nat. Nanotechnol.*, vol. 3, pp. 201–205, 2008.

[69] R. Kim and M. S. Lundstrom, "Characteristic Features of 1-D Ballistic Transport in Nanowire MOSFETs," *IEEE Trans. Electron Devices*, vol. 7, pp. 787–794, 2008.

[70] S. O. Koswatta, S. Hasan, M. S. Lundstrom, M. P. Anantram, and D. E. Nikonov, "Ballisticity of nanotube field-effect transistors: Role of phonon energy and gate bias," *Appl. Phys. Lett.*, vol. 89, no. 2, p. 023125, 2006.

[71] K. Natori, Y. Kimura, and T. Shimizu, "Characteristics of a carbon nanotube field-effect transistor analyzed as a ballistic nanowire field-effect transistor," *J. Appl. Phys.*, vol. 97, no. 3, p. 034306, 2005.

[72] A. Razavieh, D. B. Janes, and J. Appenzeller, "Transconductance Linearity Analysis of 1-D , Nanowire FETs in the Quantum Capacitance Limit," *IEEE Trans. Electron Devices*, vol. 60, no. 6, pp. 2071–2076, 2013.

[73] A. U. Alam, C. Martin, S. Rogers, N. Paydavosi, K. D. Holland, and S. Ahmed, "RF Linearity Potential of Carbon-Nanotube Transistors Versus MOSFETs," *IEEE Trans. Electron Devices*, vol. 12, no. 3, pp. 340–351, 2013.

[74] X. Zhou, J.-Y. Park, S. Huang, J. Liu, and P. McEuen, "Band Structure, Phonon Scattering, and the Performance Limit of Single-Walled Carbon Nanotube Transistors," *Phys. Rev. Lett.*, vol. 95, no. 14, p. 146805, Sep. 2005.

[75] J. Guo and M. Lundstrom, "Role of phonon scattering in carbon nanotube field-effect transistors," Appl. Phys. Lett., vol. 86, no. 19, p. 193103, 2005.

[76] O. Balci and C. Kocabas, "High frequency performance of individual and arrays of single-walled carbon nanotubes," Nanoscale, vol. 23, p. 245202, 2012.

[77] V. Perebeinos, J. Tersoff, and P. Avouris, "Electron-Phonon Interaction and Transport in Semiconducting Carbon Nanotubes," *Phys. Rev. Lett.*, vol. 94, no. 8, p. 086802, Mar. 2005.

[78] Y.-F. Chen and M. Fuhrer, "Electric-Field-Dependent Charge-Carrier Velocity in Semiconducting Carbon Nanotubes," *Phys. Rev. Lett.*, vol. 95, no. 23, p. 236803, Nov. 2005.

[79] S. Salamat, S. Member, X. Ho, J. A. Rogers, M. A. Alam, A. The, and I. V, "Intrinsic Performance Variability in Aligned Array CNFETs," *IEEE Trans. Electron Devices*, vol. 10, no. 3, pp. 439–444, 2011.

[80] A. E. Islam, F. Du, X. Ho, S. H. Jin, and S. Dunham, "Effect of variations in diameter and density on the statistics of aligned array carbon-nanotube field effect transistors," *Appl. Phys. Lett.*, vol. 111, p. 054511, 2012.

[81] W. Zhou, L. Ding, S. Yang, and J. Liu, "Synthesis of High-Density Large-Diameter, and Aligned Single-Walled Carbon Nanotubes by Multiple-Cycle Growth Methods," *ACS Nano*, vol. 5, pp. 3849–3857, 2011.

[82] W.H. Chiang and R. M. Sankaran, "Linking catalyst composition to chirality distributions of as-grown single-walled carbon nanotubes by tuning Ni(x)Fe(1-x) nanoparticles.," *Nat. Mater.*, vol. 8, no. 11, pp. 882–6, Nov. 2009.

[83] D. M. N. M. Dissanayake and Z. Zhong, "Schottky diodes using as-grown single-walled carbon nanotube ensembles," *Appl. Phys. Lett.*, vol. 104, no. 12, p. 123501, Mar. 2014.