Fabrication of Carbon Nanotube Field Effect Transistor Using Dielectrophoresis and Its Application as Static Random Access Memory Bit Cell

Shobhit Kareer

A thesis submitted in conformity with the requirements for the degree of Master of Applied Science in Electrical and Computer Engineering

Ottawa-Carleton Institute for Electrical and Computer Engineering

Faculty of Engineering

University of Ottawa

© Shobhit Kareer, Ottawa, Canada, 2019
The undersigned hereby recommends to the
Faculty of Graduate and Postdoctoral Affairs
acceptance of the thesis

Fabrication of Carbon Nanotube Field Effect Transistor
Using Dielectrophoresis and its Application as Static Random Access Memory Bit cell

submitted by Shobhit Kareer
in partial fulfilment of the requirements for the degree of
Master of Applied Science in Electrical and Computer Engineering

Professor Jeongwon Park, Thesis Supervisor
School of Electrical Engineering and Computer Science

Professor Voicu Groza, Examiner
School of Electrical Engineering and Computer Science
University of Ottawa

Professor Steve McGarry, Examiner
Department of Electronics
Carleton University

Professor Guy-Vincent Jourdan, Chair
School of Electrical Engineering and Computer Science
University of Ottawa

Ottawa-Carleton Institute for Electrical and Computer Engineering
School of Electrical Engineering and Computer Science
University of Ottawa, August 2019
Abstract

The aim of the thesis is to fabricate Schottky contact carbon nanotube field effect transistor (CNFET) using the dielectrophoresis (DEP) to resolve the alignment issue and show its transistor behaviour. The work presented is a combination of fabrication and simulation of CNFET. Fabrication of the device electrode had been done using the electron beam lithography to achieve a channel length of 150nm and analysis was done on an optical microscope, SEM, AFM and Raman spectroscopy. Second half of the thesis provides a solution to “bottleneck communication” between microprocessor and memory to increase the computation for applications like AI, IoT etc and 3D monolithic memories. As a solution, we propose a novel CNFET based processing in-memory architecture using a novel CNFET dual port single-ended SRAM bit cell. The combination of the CNFET and processing in-memory can be a new phase for memory and computation.
Acknowledgement

I would like to take this opportunity to express gratitude towards my supervisor, Prof. Jeongwon Park, for his valuable guidance, experience and inspiration; thank you for being more than a thesis supervisor. A special thanks to Centre for Research in Photonics’ (CRPuO) director, Prof. Pierre Berini for his support and his lab technologist, Anthony Olivieri for his time and training on the tools offered by the CRPuO.

A special acknowledgment to Stanford University's Professor, Prof. H.-S. Phillip Wong and his student, Yujo for providing us with the wafer and guidance for the project.

I would also like to thank my God for giving me patience and express my gratitude towards my parents and sisters for supporting me financially and emotionally in my journey towards my dream.

I would like to thank Prof. Voicu Groza and Prof. Steve McGarry for taking out the time for my thesis review.

I would also like to acknowledge Center for Advance Material Research (CAMaR) facility for providing Raman spectroscopy and finally, CMC Microsystems for the provision of CAD tools (Cadence IC 6.17) and equipment loan (probe station, Keighley source and LCR meter) that facilitated this research simulations and characterizations.
# Table of Content

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abstract</td>
<td>iii</td>
</tr>
<tr>
<td>Acknowledgement</td>
<td>iv</td>
</tr>
<tr>
<td>List of Table</td>
<td>xi</td>
</tr>
<tr>
<td>List of Acronym</td>
<td>xii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Motivation</td>
<td>2</td>
</tr>
<tr>
<td>About this thesis</td>
<td>2</td>
</tr>
<tr>
<td>Chapter 1: Carbon Nanotube Structure Synthesis, Structure and Characterization</td>
<td>4</td>
</tr>
<tr>
<td>1.1 Structure of CNTs</td>
<td>4</td>
</tr>
<tr>
<td>1.1.1 Single Wall or Multi Wall Carbon Nanotube</td>
<td>4</td>
</tr>
<tr>
<td>1.2 Synthesis of CNTs</td>
<td>5</td>
</tr>
<tr>
<td>1.2.1 Arc Discharge Method</td>
<td>6</td>
</tr>
<tr>
<td>1.2.2 Laser Ablation</td>
<td>6</td>
</tr>
<tr>
<td>1.2.3 Chemical Vapor Deposition</td>
<td>7</td>
</tr>
<tr>
<td>1.3 Characterization Techniques</td>
<td>8</td>
</tr>
<tr>
<td>Chapter 2: Nano Carbon Structure Properties</td>
<td>10</td>
</tr>
<tr>
<td>2.1 Overview of graphene</td>
<td>10</td>
</tr>
<tr>
<td>2.1.1 Direct Lattice</td>
<td>10</td>
</tr>
<tr>
<td>2.1.2 Tight Binding dispersion</td>
<td>10</td>
</tr>
<tr>
<td>2.2 Carbon Nanotube</td>
<td>11</td>
</tr>
<tr>
<td>2.2.1 Chirality</td>
<td>12</td>
</tr>
<tr>
<td>2.2.2 Lattice Structure</td>
<td>12</td>
</tr>
<tr>
<td>2.2.3 Band Structure</td>
<td>13</td>
</tr>
<tr>
<td>Chapter 3: Carbon Nanotube Field Effect Transistor</td>
<td>16</td>
</tr>
<tr>
<td>3.1 Review of CNFET structure</td>
<td>16</td>
</tr>
<tr>
<td>3.2 Schottky Barrier Ballistic CNFET</td>
<td>17</td>
</tr>
<tr>
<td>3.3 Process Variation</td>
<td>21</td>
</tr>
<tr>
<td>3.3.1 Width and Length of Carbon Nanotube Transistor</td>
<td>22</td>
</tr>
<tr>
<td>3.3.2 Chirality</td>
<td>23</td>
</tr>
<tr>
<td>3.4 Introduction to Dielectrophoresis</td>
<td>24</td>
</tr>
<tr>
<td>Summary</td>
<td>26</td>
</tr>
<tr>
<td>Chapter 4: Experimental Process and Characterization</td>
<td>27</td>
</tr>
</tbody>
</table>
4.1 Fabrication Process Flow 27
  4.1.1 Lithography 28
  4.1.2 Metal Deposition 28
4.2 Solution Preparation of CNTs 28
4.3 Electrode Design 29
4.4 Dielectrophoresis 31
4.5 Annealing 33
4.6 Characterization 33

Conclusion 37

Chapter 5: Application as a Memory: Static Random Access Memory 39
  5.1 Introduction 39
  5.2 Static Random Access Memory 40
  5.3 CNFET based Single Ended SRAM with Processing-in-Memory 42
    5.3.1 Proposed SRAM bit cell 48
    5.3.2 CNFET SRAM Bit Cell Performance 51
    5.3.3 Layout 56
    5.3.4 CNFET based in-memory computing 58

Conclusion 64

Future Work 65
References 66
Appendix 72
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Monotonous increase in the gate length which is in contrast to the predicted scaling [1] © 2016 IEEE</td>
</tr>
<tr>
<td>3</td>
<td>First reported TEM images of MWCNTs, the upper and lower MWCNT have diameter of 50nm and 100nm respectively [16]</td>
</tr>
<tr>
<td>4</td>
<td>Arc Discharge synthesis schematic representation [17]; (b) SEM image of the Arc Discharge method synthesized by Raymor Industries, Quebec.</td>
</tr>
<tr>
<td>5</td>
<td>Schematic diagram of Laser Ablation CNT synthesis [19]</td>
</tr>
<tr>
<td>6</td>
<td>(a) Schematic representation of CVD (b) Growth mechanism in CVD</td>
</tr>
<tr>
<td>7</td>
<td>Raman Analysis done by the Raymore Industries for the batch of CNTs used in the work</td>
</tr>
<tr>
<td>8</td>
<td>TEM image adapted from [23] to show the lattice structure of the graphene</td>
</tr>
<tr>
<td>9</td>
<td>MATLAB simulation of band structure of graphene (a) 2-D, (b) 3-D and (c) Contour plot of one band showing the lattice structure of graphene</td>
</tr>
<tr>
<td>10</td>
<td>Representation of Chirality and unrolled various types of Single Wall CNTs</td>
</tr>
<tr>
<td>11</td>
<td>MATLAB simulation of band structure of CNTs having chirality (a) Armchair CNT (9,0) and (b) Zig-zag CNT (10,0)</td>
</tr>
<tr>
<td>12</td>
<td>Band structure of CNT, the shift represented by dash line is due to an applied voltage over the CNT. (For simplicity the sub-bands of the CNTs are not shown)</td>
</tr>
<tr>
<td>13</td>
<td>Different configuration for Carbon Nanotube Field Effect Transistor</td>
</tr>
<tr>
<td>14</td>
<td>Schottky behaviour between CNT and metal (Gold and Palladium)</td>
</tr>
<tr>
<td>15</td>
<td>Band diagram of n-CNFET with Schottky barrier, where Ø represents the Schottky barrier height for the holes to transfer from the valence band of the source to the conduction band of the drain (a) gate voltage is zero; (b) gate voltage at linear regime ( V_D = V_g - V_t ) . Adapted from [2]. Here ( E_f ) is the fermi level, ( V_D ) is the drain voltage, ( V_g ) is the gate voltage and ( V_t ) is the threshold voltage.</td>
</tr>
</tbody>
</table>
Figure 16 Conduction band profile for Schottky barrier in n-CNFET. Adapted from [2].

Figure 16: (a) Simulation of drain characteristic of a CNFET with different channel gate length \( L_g \) (b) Simulation of drain characteristics with different number of carbon nanotubes in a CNFET

Figure 17: 6σ Monte Carlo variation of chiral indexes for 10 points using Gaussian distribution for chiral vector (19,0)

Figure 18: Circuit diagram for implementing capacitive coupled DEP.

Figure 19: Overview of the fabrication process used in this work

Figure 20: (a) SEM image of the electrode fabricated using e-beam lithography (b) Simulation for the electric field intensity around the fabricated electrode at 50 kHz and also showing the row series and parallel rows CNFET configuration.

Figure 21 Electric Field formation around the CNTs at (a) 50kHz (b) 150kHz

Figure 22 SEM image of the Transistor using the DEP (a) at 50Hz, 2V_{PP} and 5 secs (b) at 150kHz, 5V_{PP} and 5 Secs

Figure 23: (a) Pictorial Representation of the characterization setup (b) Device under test (DUT) snapshot

Figure 24: (a) Drain characteristics with the gate voltage (5,10, and 15V) showing the n-type CNFET, before annealing (b) Shows the drain characteristics after annealing at 350° C.

Figure 25: (a) Drain characteristics with the gate voltage (5,10, and 15V) showing the p-type CNFET, before annealing (b) Shows the drain characteristics after annealing at 350° C

Figure 26: Simulation and experimental drain characteristics of the fabricated CNFET.

Figure 27: Flowchart summary for the DEP process

Figure 28 Conventional 6T SRAM bit cell architecture

Figure 29 Conventional 6T SRAM bit cell architecture

Figure 30: Voltage Transfer Characteristic (VTC) of CMOS and CNT based inverter and Schmitt Trigger Inverter

Figure 31: CNFET model used for simulation which is based on top gate formation [71].
Figure 32: Hysteresis loop formed by the Schmitt trigger used in the proposed SRAM bit cell

Figure 33: Conventional (a) 6T SRAM bit cell (b) 8T SRAM bit cell

Figure 34: Von Neumann Architecture showing the bottleneck of data transfer between the CPU and memory.

Figure 35: Proposed SRAM bit cell

Figure 36: Write operation with varying BL voltage to show the respective transient signal integrity

Figure 37: (a) WSNM test configuration using the word line sweep method (b) WSNM comparison between comparison and conventional SRAM bit cells.

Figure 38: Comparison of Read SNM between proposed SRAM bit cell and conventional SRAM bit cells

Figure 39: Read SNM simulation configuration

Figure 40: Performance comparison of the CNFET SRAM bit cells

Figure 41: 1000 points Monte-carlo simulation of Read-SNM for the proposed SRAM bit cell

Figure 42: Write time v/s density distribution for 6 σ variation for the proposed memory cell

Figure 43: (a) Average Power distribution and (b) Hold Power distribution of the 8T conv., 6T conv. and Proposed SRAM bit cell over the 6σ density variation for 1000 random Gaussian distributed sample

Figure 44: Mask 1 for creating DEP for fabricating the proposed SRAM

Figure 45: Final layout for the proposed SRAM bit cell

Figure 46: (a) Proposed CNFET SRAM in-memory computing architecture, (b) Circuit used for the logical in-memory computation for the proposed SRAM bit cell.

Figure 47: NOR operation using the proposed SRAM using the in-memory computing configuration

Figure 48: NAND operation using the proposed SRAM using the in-memory computing configuration

Figure 49: XOR operation using the proposed SRAM using the in-memory computing configuration
Figure 50: Raman spectroscopy for the CNTs on the sample over 4 spots using 735nm near IR laser, the area between the red line shows the presence of metallic-CNTs.

Figure 51: Raymore CNT in Toluene

Figure 52: CNTs in (Iso-propanol Alcohol) IPA

Figure 53: Capacitance Measurement of the silicon wafer.

Figure 54: Refractive index of the oxide used v/s the incident wavelength using Ellipsometry

Figure 55: SEM for the electrode with DEP parameters $2 V_{pp}, 150 \text{ kHz and 6 secs}$

Figure 56: SEM for the electrode with DEP parameters $2 V_{pp}, 150 \text{ kHz and 10 secs}$

Figure 57: AFM for the electrode with DEP parameters $2 V_{pp}, 150 \text{ kHz and 3 secs}$

Figure 58: AFM for the electrode with DEP parameters $2 V_{pp}, 150 \text{ kHz and 10 secs}$

Figure 59 AFM for the electrode with DEP parameters $2 V_{pp}, 150 \text{ kHz and 20 secs}$
List of Table

Table 1 Parameters used for simulating CNTs in HFSS 32
Table 2: Truth table for NOR gate 59
Table 3: Truth table for NAND gate 61
Table 4: Truth table for XOR gate 62
Table 5 Resist layer 1 recipe 73
Table 6 Resist layer 2 recipe 74
List of Acronym

- Artificial Intelligence  
  AI
- Atomic force microscope  
  AFM
- Back Gate  
  BG
- Carbon nanotube  
  CNT
- Carbon nanotube field effect transistor  
  CNFET
- Complementary metal oxide semiconductor  
  CMOS
- Computation random access memory  
  CRAM
- Chemical vapour deposition  
  CVD
- Dielectrophoresis  
  DEP
- Extreme ultraviolet  
  EUV
- Field effect transistor  
  FET
- Internet of things  
  IoT
- Multi wall carbon nanotube  
  MWCNT
- Processing in-memory  
  PIM
- Read static noise margin  
  RSNM
- Scanning electron microscope  
  SEM
- Static random access memory  
  SRAM
- Single wall carbon nanotube  
  SWCNT
- Top gate  
  TG
- Write static noise margin  
  WSNM
Introduction

Motivation

With a simple remark that transistors which are integrated on semiconductor circuit are going to scale up twice every 18 months [3], Gordon Moore placed the foundation for a race towards reaching the smallest feature size for a transistor.

The advances in circuit complexity categorized by him are based on three principle technical drivers which are increasing chip density, decreasing feature size and improving signal processing. Though achieving aggressive scaling for silicon surfaces is now possible with techniques like extreme UV (EUV) lithography, but it comes with a price of degradation in signal integrity and process complexity at sub-micrometer regime [4]. Therefore from figure 1, we can notice stagnation of gate length that indicates the limitation of scaling in silicon technology.

![Image](image.png)

Figure 1: Monotonous increase in the gate length which is in contrast to the predicted scaling [1] © 2016 IEEE

So to achieve a noise resilient signal processing and robust memory which are capable for AI and Internet of things (IoT)’s applications, hardware designer are now exploring new material technology and architecture to design electronics based on uprising material, often categorized as beyond CMOS materials, for example carbon [3], III-V semiconductors[5], phase change materials [6], etc.
However, the quasi one directional Carbon Nanotubes (CNTs) is a near-future material for replacing silicon channel in the transistors, also known as Carbon Nanotube Field Effect Transistor (CNFET). This potential material shows promising properties in thermal conductivity [8], electrical transport [9] and optical transmittance [10]. According to International Technology Roadmap for Semiconductors (ITRS), shown in figure 2, the projected application of CNTs and graphene devices is the most promising advancement for Metal Oxide Semiconductor (MOS) industry which can lead to new logic designing and architecture.

Along with the characteristics mentioned earlier, CNFETs are considered promising because of its adaptability to the existing logic architecture based on CMOS technology. Moreover, CNFET application promises a new alternative to the thin-film transistor (TFT), flexible electronics, radio frequency and sensor applications [7].

With the invention of CNTs, it has gained a lot of popularity with its one-dimensional structure. However, manufacturing of CNTs is a stochastic process and depositing them at the desired location on a surface is an arduous task, therefore CNTs have so far not seen success in commercial electronics. As the process of fabrication for CNFET would be tough with a low yield (within the wafer). Therefore, it motivated me to find a solution for making a high yielding VLSI compatible CNT placement solution which can be implemented in a layout for a memory architecture.

**About this thesis**

“Fabrication of Carbon Nanotube Field Effect transistor Using Dielectrophoresis and Its Application as Static Random Access Memory Bit Cell”, this thesis can be broadly be divided into 3 parts: theory, experimental and simulation.
Chapters 1, 2 and 3 are the theoretical and background research for the carbon nanotube and carbon nanotube field effect transistor. Chapter 1 and 2 discuss the types of carbon nanotube, how they can be synthesized, characterization techniques and basic physics behind the carbon nanotubes. Subsequently, Chapter 3 discusses the CNFET’s overview on the types of structures, Schottky contact in CNFET, why we chose palladium over gold for the contact, the parameters to be considered while simulating a CNFET based circuit’s process variation and theory of dielectrophoresis (DEP).

Chapter 4 consists of the experimental part which discusses in detail about the process followed for making a back-gate Schottky contact CNFET using the DEP, along with the finite element analysis (FEA) for the DEP using ANSYS HFSS to show why 50 kHz was chosen for the deposition. To finish, chapter 4 shows the characteristic setup and the drain characteristics for the fabricated transistor.

Chapter 5 discusses the application of CNTs as static random access memory (SRAM) bit cell using the Stanford simulation model. The introduction to SRAM bit cell and in-memory computing is given using the CMOS technology and it’s further improvement by implementing it on the CNFET technology. Also, the chapter showcases a novel architecture for the SRAM bit cell. SRAM characteristics are simulated along with its process variation using Verilog-A in Cadence Spectre to account for the stochastic characteristic variation of the CNTs that can be produced due to process or mismatch variations. Secondly, we show on paper how DEP can be implemented on the layout for the proposed SRAM bit cell, and lastly, we exhibited how the proposed CNFET SRAM bit cell can be useful for implementing a non-von Neumann processing in-memory architecture.

Chapter 1: Carbon Nanotube Structure Synthesis, Structure and Characterization

Carbon is the most captivating material in nature; it is found in environment in many allotropic forms, they exists on Earth both as an expensive as well as a cheap material, i.e. diamond and graphite, respectively. Carbon allotropes have been extensively used and explored for centuries in various products, nonetheless, new crystalline forms of carbon have been recently discovered which includes Buckminsterfullerene (C60 or Buckyball), carbon nanotube and graphene. These allotropes show distinct properties than the previous allotropes. They all have a hexagonal lattice of carbon atoms in common besides their nanometre dimensions.

In this work, CNTs and graphene has been studied. But the primary focus is on CNTs because of its high conductivity, near ballistic nature and high tensile strength; which poses a potential uses in 3D IC that can have a major impact on computer architecture and performance [11], thus pushing the limits of Moore’s Law.

Since 1991, the discovery of CNTs by Sumio Ijima [12], many diversified applications and properties have been researched. Synthesis of CNTs has been advanced and made more profuse. Now, it can be grown up to millimetre in length and with a range of diameters. These intriguing properties of CNT give a range of options for various applications like interconnects [13], atomic force microscope probe tips [14] or ballistic field effect transistor (FET) [15]. All of these are dependent on the structure of CNTs, which is described briefly below.

1.1 Structure of CNTs

CNTs unique properties correspond to its extraordinary structure which leads to various applications. These quantum wires can have a range of diameter from 0.8nm to 45nm with length reaching the thousands of micrometres. These nanowires are the results of dangling bond in graphene sheets trying to lower the large energy forming a tubular structure. Thus the CNTs’ structural and lattice characteristics can be associated with graphene.

1.1.1 Single Wall or Multi Wall Carbon Nanotube

The cylindrical CNTs can be grown broadly in two forms: single wall CNT (SWCNT) or multi wall CNT (MWCNT). MWCNT consists of concentric SWCNTs cylinder separated by the van der waals distance, which is approximately 3.4Å [16].
In 1952, Radushkevich et al. reported the first transmission electron microscope (TEM) images of MWCNT (Figure 3). This image also starts a controversy whether the CNTs were found in 1991 or in 1952.

The concentric shells of MWCNT are not always uniform and can differ in nature. For example, the outer cylinder can be metallic whereas the inner cylinder can be semiconducting. Therefore, it becomes a challenge to use MWCNTs in the structure because if the outer shell is metallic and the inner shell is semiconducting the outer shell can suppress the semiconducting nature of the inner shell. One of the ways to dissipate such defect is called electric annealing, in which large current is applied across the MWCNTs to burn off the metallic CNTs [16]. However, the disadvantage of this process is that the outer shell of the CNTs are eliminated first due to direct contact with the electrode followed by the inner CNTs. Moreover, it can damage the inner semiconducting CNTs as well because of thermal accumulation [16].

Single wall carbon nanotube (SWCNT) is made of cylindrical graphene which can be either metallic or semiconducting in nature. Since the SWCTs are cylindrical, its properties are related to the tube structure, which has been discussed in detail in the following chapter. Because the CNT used in the work is SWCNT, only SWCNTs is discussed and further on will be described as CNTs instead of SWCNTs.

Since, the structure of CNTs plays a role in defining whether the CNTs are metallic or semiconducting, it becomes essential to choose an appropriate option of synthesizing which is producing the desired CNTs’ structure and are cost-effective as well.

1.2 Synthesis of CNTs

Due to its potential market in various application CNTs’ synthesis can be done using various methods, but the most prevailing methods are arc discharge, laser ablation and chemical vapour deposition (CVD).
1.2.1 Arc Discharge Method

![Arc Discharge Method Diagram]

Figure 4: Arc Discharge synthesis schematic representation [17]; (b) SEM image of the Arc Discharge method synthesized by Raymor Industries, Quebec.

Applying a very high electric voltage between two graphite rods under high temperature leads to the formation of an electric arc of carbon atoms. This carbon plasma is attracted to the cathode. The plasma re-condenses as it approaches the metal catalyst, like iron, nickel etc which leads to the growth of CNTs. Since the method is an exothermic reaction, the synthesis is done in a water-cooled helium chamber at sub-atmospheric pressure, which plays an important role in the structure formation of the carbon. Figure 4 shows the schematic representation of the CNTs synthesis for the arc discharge method and the SEM image of arc discharged CNTs used in this work. The separation of the graphite electrode is in range of millimeter and the voltage across the arc is in range of 20V. The quantity produced during this process is less than 40% by weight, but the process creates a mixture of MCNT and SWCNTs with an almost defect-free structure and length less than 50µm [18]. The main disadvantage of this synthesis is the random placement of the CNTs for an application.

1.2.2 Laser Ablation

This synthesis is based on the removal of matter from carbon based surface with the application of laser, hence it is known as ablation (which means the removal of matter). In this method, the surface of graphite is heated to high temperature in a reactor which is targeted by the laser, the graphite surface contains CNT catalyst metal like nickel-cobalt, to increase the quantity of CNTs.

Figure 5, shows the schematic representation of the laser ablation CNT synthesis, which depicts the graphite source targeted by a laser inside a tubular furnace and is being collected at other end made of copper (Just like in arc discharge method).
The laser ablation is an advancement to the arc discharge method as it gives better control over the quantity and purity of the CNTs. It is observed that the CNTs structure can be controlled by regulating the interval of interaction between graphite and laser [19]. But these CNTs is more expensive than arc discharged CNTs.

1.2.3 Chemical Vapor Deposition

Since 1960 CVD has been used for synthesizing carbon filament, but now the CVD can be used for CNT growth as well. This method comprises of the decomposition of carbon rich gas onto a substrate in the presence of a metal catalyst at a high temperature (∼600°C). Figure 6 (a) shows the schematic for CVD and Figure 6 (b) shows the growth of CNT as a process of nucleation of the catalyst and decomposition of the precursor. It is one of the cheapest options for growing CNTs and it gives control of orientation over the substrate. But the CNTs grown using this method have defects [20].
1.3 Characterization Techniques

Characterization of CNTs requires experimental skills in order to determine the CNT physical dimension, structure quality, placement and electronic characteristics. In this work, routine characterization techniques: scanning electron microscope (SEM), atomic force microscope (AFM) and Raman spectroscopy have been used for characterization. Raman spectroscopy has been selected to provide comparatively faster confirmation of the presence of CNTs on the substrate, whereas the SEM and AFM gives visual image of the CNTs.

Raman spectroscopy gives electronic characterization of CNTs by probing CNT lattice which gives distinct characterization peaks based on the atomic structure and bonds present on CNTs. These characteristics have been performed by the Raymore Industries, and has been taken as reference in this work. We have used their Super Pure CNTs. In addition, Radial breathing mode (RBM) phonon can be used for the chirality characterization of the CNTs. In this mode, atoms of the carbon structure oscillate radially with the same phase generating a strong peak around 300 cm\(^{-1}\) in Raman Spectroscopy. This can be used for getting an average number for the CNT’s diameter, given by equation 1.

![Figure 7: Raman Analysis done by the Raymore Industries for the batch of CNTs used in the work](image)

The relation between the Raman Spectroscopy and CNT’s diameter is given by the formula

\[ \text{Wave number} = \frac{248}{d} \, \text{cm}^{-1} \]

equation 1

Where \(d\) is the tube diameter in nm.

G band peak shown in Figure 7, illustrates metallic content present in the CNT thick film produced by the Raymore Industries. The sharp peak at 1600 cm\(^{-1}\) shows the semiconducting
CNTs. Whereas, broadening of the peak between 1450 cm$^{-1}$ - 1550 cm$^{-1}$ shows metallic CNTs present in the batch [21]. Additionally, there is a discrepancy in the data provided by the Raymore Industries and ours (Appendix: Figure 51), because according to our Raman spectrum data we had a significant D peak that correspond to structural damage of CNTs [22].
Chapter 2: Nano Carbon Structure Properties

2.1 Overview of graphene

Carbon in the elementary state has four valence electrons in its 2s and 2p orbitals and can have various allotropes based on its hybridization. Graphene is a planar allotrope formed from the sp\(^2\) hybridization. In graphene 2s orbital interacts with 2p\(_x\) and 2p\(_y\) orbitals to form the hybridization constructing three \(\sigma\)-bonds along the planar axis. These \(\sigma\)-bonded carbon atoms are responsible for great strength and mechanical properties for graphene. Meanwhile the p orbital electrons are weakly bonded and delocalized, resulting in the electronic properties of graphene and CNTs.

2.1.1 Direct Lattice

Graphene has a honeycomb lattice as shown in Figure 8, where the gray dot represents a carbon atom with carbon-carbon bond length (\(a_{c-c}\)) which is approximately 1.42 Å. The honeycomb structure is the Bravais lattice of graphene which has a basis of 2 atoms, shown in the Figure 8 as A and B. These atoms contribute two \(\pi\) electrons per unit cell for the electronic properties of graphene.

![Figure 8: TEM image adapted from [23] to show the lattice structure of the graphene](image)

2.1.2 Tight Binding dispersion

The Bravais lattice described in previous section are defined as

\[ A = \frac{\sqrt{3}}{2} a\hat{x} + \frac{1}{2} a\hat{y} \]  
\[ B = \frac{\sqrt{3}}{2} a\hat{x} - \frac{1}{2} a\hat{y} \]

Where \(a = \sqrt{3}a_{c-c} = 2.46\text{Å}\), and \(a_{c-c}\) is the length of carbon-carbon bond.
Figure 9: MATLAB simulation of band structure of graphene (a) 2-D, (b) 3-D and (c) Contour plot of one band showing the lattice structure of graphene

Since graphene has two basis, the wave function of graphene is addition of two sub-lattice Bloch function:

$$\Psi(k,r) = C_A \varphi_A(k,r) + C_B \varphi_B(k,r) \quad \text{equation 4}$$

Where $r$ is the lattice vector and $\varphi$ is the Wannier function that defines the approximation of linear combination of atomic orbitals in k-space when considered in a specific area to avoid distortion of the atomic orbital. Since graphene is sp$^2$ hybridized, the electrons shared within these bonds are considered in the calculation. Therefore, the potential term is given by the summation of every atom’s potential energy.

The band energy dispersion for the nearest neighbor model is given by the equation 5 and is shown in Figure 9.

$$E(k) = \pm \gamma |\zeta(k)| = \gamma \sqrt{1 + 4 \cos \left(\frac{\sqrt{3}ak_y}{2}\right) \cos \left(\frac{ak_y}{2}\right) + 4 \cos^2 \left(\frac{ak_y}{2}\right)} \quad \text{equation 5}$$

Where $\gamma$ is the nearest neighbor constant and $\zeta$ is the interference factor

### 2.2 Carbon Nanotube

With diameter in the range of 0.5nm to 5nm, carbon nanotube are the quantum wires formed by rolling up of graphene sheet. The two basic forms of CNTs are SWCNTs and MWCNTs, which has been discussed earlier in chapter 1. In this section, the CNTs has been further characterized based on chirality, lattice structure and band structure.
2.2.1 Chirality

As mentioned in Section 1.1, CNTs can configure itself as semiconducting and metallic CNTs, which depends on the key characteristic of CNTs known as chirality. Chirality can also be described as a wrapping of monolayer graphene along two points, known as chiral vector, to form a tube. Chiral vector can mathematically be expressed as,

\[ C_v = n_1 \hat{a}_1 + n_2 \hat{a}_2 \]  

where \( \hat{a} \) is the lattice unit vector separated by an angle of 60\(^\circ\) forming the CNTs’ Bravais lattice and, \( n_1 \) and \( n_2 \) are the positive indices (\( 0 \leq n_1 \leq n_2 \)), shown in Figure 10.

Depending on the rolling of graphene sheet, CNTs can be categorized as armchair, zig-zag and chiral CNTs. These CNTs are classified based on the edges of the tube. Figure 10 shows the armchair CNT edge in red and zig-zag CNT edge in blue. Additionally, CNTs can also be categorized based on the chiral vector \((n_1, n_2)\), where \( n_1 \) and \( n_2 \) are integer value. These vectors determines the nature of the chiral CNTs. That is, if \( n_1 = n_2 = 3i \), where \( i \) is an integer, then nanotube will be in an armchair formation and forms a metal CNT [24].

2.2.2 Lattice Structure

Enduring the consideration that CNTs are tubular form of the graphene sheet, which is rolled along the chiral vector \((n_1, n_2)\), the primitive lattice structure formation of CNT becomes easier. Now using the chiral vector, we can derive the circumference of the CNT \( (C_h) \) as well as chiral angle, which is calculated as:

\[ C_h = a \sqrt{n_1^2 + n_2^2 + n_1n_2} \]  

\[
\cos \theta = \frac{c_h \cdot a_1}{|c_h||a_1|} = \frac{2n_1 + n_2}{2\sqrt{n_1^2 + n_2^2 + n_1 n_2}} \quad \text{equation 8}
\]

where, \( \theta \) is the chiral angle.

Also, chirality plays an important role in determining the diameter of the CNT, which is given by the following formula

\[
D_{CNT} = \frac{|c_h|}{\pi} = \frac{\sqrt{n_1^2 + n_2^2 + n_1 n_2}}{\pi} \quad \text{equation 9}
\]

Where \( D_{CNT} \) is the diameter of the CNT. Different nanotubes can have same diameter but different chirality. As a result, diameter cannot be a unique parameter to characterize a batch of CNTs [25].

### 2.2.3. Band Structure

CNTs’ band structure can be determined using the near neighbor tight binding (NNTB) energy dispersion model of graphene, which is mentioned in earlier sub-section. Since wrapping of the CNT have same effect on the Brillouin zone of graphene as overlapping of any primitive unit cell. However, this is not valid for small diameter \( D_{CNT} < 1 \text{nm} \) or high fermi level energy CNT [26]. Thus, using the equation 5, and using the chiral vector we can find the band structure of the CNTs which are dependent on the type of chirality the CNTs have. For the zig-zag CNTs energy dispersion-\( k \) vector space equation is given by:

\[
E(k) = \pm \gamma \sqrt{1 + 4 \cos \left( \frac{\sqrt{3} a k_y}{2} \right) \cos \left( \frac{\pi j}{n_1} \right) + 4 \cos^2 \left( \frac{\pi j}{n_1} \right)} \quad \text{equation 10}
\]

and armchair CNT \( E-k \) dispersion is given by:
\[ E(k) = \pm \gamma \sqrt{1 + 4\cos(\frac{\sqrt{3}ak}{n_1})\cos(\frac{n_1}{2}) + 4\cos^2(\frac{\pi j}{2})} \quad \text{equation 11} \]

where, \( \gamma \) is the nearest neighbor overlap constant (3.1 eV), \( j \) is the sub-band integer index (1 to \( 2n_1 \)), and \( a \) is the graphene Bravais lattice constant (2.46 Å). From section 2.2.1 and Figure 11 (a), we can say that all the armchair CNTs are semi-metal. Whereas, other CNTs have a band gap between the conduction and valence band, thus making them semiconductor (Figure 11 (b)).

Band gap is important while designing an electronic component. In CNTs, band index of the first subband (\( j_1 \)) for semiconducting CNT is closest integer to \( 2n_1/3 \).

The energy band in a CNT is given by

\[ E_g = 2\gamma \frac{a_{c-c}}{D_t} \quad \text{equation 12} \]

where \( E_g \) is the bandgap energy, \( \gamma \) is the nearest neighbor constant, \( a_{c-c} \) is the distance between two carbon atoms in the lattice (1.42 Å) and \( D_t \) is the CNT diameter.

\[ \text{Figure 12: Band structure of CNT, the shift represented by dash line is due to an applied voltage over the CNT. (For simplicity the sub-bands of the CNTs are not shown)} \]

Band structure of CNTs gives us the surface potential, which is defined as the difference between intrinsic fermi level and fermi level at the surface [27]. It accounts for the excess charge induced in the semiconductor due to an external field. A simple band structure of CNTs is shown in Figure 12. Due to an applied field more than 0 eV, conduction band can be shifted near to fermi level, thus the CNT is now n-type CNTs. Whereas, if the potential is less than zero, the conduction band moves away from the fermi level and therefore act as a p-type CNT. Following the above statement, we can say that the CNTs are ambipolar in behavior, that is, it
can act both as an n-type and as a p-type material depending on the potential applied. Surface potential of CNT in the transistor formation is determined by the capacitive coupling of the gate to the CNT, which is given by the formula [28]

\[ \phi_s = \frac{C_{\text{oxide}}}{C_{\text{oxide}} + C_q} V_g \]  

*equation 13*

where \( C_{\text{oxide}} \) is capacitance created by the transistor’s oxide, \( C_q \) is quantum capacitance created between the CNTs and gate, and \( V_g \) is gate voltage.

When, \( \phi > 0 \) the CNTs are n-type, when \( \phi = 0 \) the CNTs are intrinsic and when \( \phi < 0 \) the CNT are p-type; this convention provides us with convenience to make an analogy between the MOSFET and CNFET.
Chapter 3: Carbon Nanotube Field Effect Transistor

As discussed briefly in the last chapter, the relative concept of a carbon nanotube as field effect transistor is similar to that of metal oxide semiconductor (MOS) field effect transistor. Analogous to that of MOS technology, CNFET also contains two terminals: source and drain which have a potential difference between them and where the flow of electrons take place. This flow of carrier is controlled by a third terminal, gate, attached to the voltage; by applying a horizontal electric field gate terminal changes the surface potential of CNFET.

![Diagram of Carbon Nanotube Field Effect Transistor](image)

*Figure 13: Different configuration for Carbon Nanotube Field Effect Transistor*

Gate can be configured into two forms: back gate (BG) and top gate (TG), shown in Figure 13. In this work we have only used the BG configuration, but the layout designed for the work is capable for having a TG.

3.1 Review of CNFET structure

On the basis of deposition of CNTs on the substrate, the CNFET structure process methodology can be classified into two categories. If the CNTs are grown on the wafer and structure (electrode) is built on them, the methodology is known as bottom up. Whereas, if the CNTs are deposited after the formation of structures the methodology is known as top down. These two methods are briefly described below using an example ech.

Top Down Methodology

In the top down approach, proposed by IBM [29], the process took on a bilayer oxide of hafnium and silicon dioxide. The CNTs used in the process were grown by arc discharge. Subsequently, segregation of metallic CNTs (m-CNTs) from semiconducting CNTs (s-CNTs) was performed in column chromatography using poly ((9,9-dioctylfluorenyl-2,7-diyl)-alt-co-(6,6′-(2,2′-bipyridine))), also known as PFO-BPy manufactured by American Dye. Following the separation, the CNT is dispersed in toluene and drop casted on the substrate having trenches of hafnium oxide bottom and walls of silicon dioxide. Before the
dispersion of the CNTs over the substrate, chloro(dimethyl)octadecylsilane is surfaced on the substrate. It acts as a self-assembly membrane (SAM) for the carbon nanotube when applied on hafnium oxide, by using the atoms of octadecylsilane which bonds the CNT and hafnium. Although the process is reliable and gives a good reproducibility, it is not yet used commercially due to the production cost of the organic polymer and hafnium oxide.

**Bottom Up Methodology**

The other method which is proposed to be compatible for VLSI design [24] and does not require drop cast of method, uses the bottom up approach by growing chemical vapour deposition (CVD) CNTs. The process gives a good CNFET in terms of alignment, but the density of the CNFET is comparatively less. Also, CNTs grown have more defects because the CNTs are grown using CVD method [24]. This top down methodology is a process of growing the CVD CNT on quartz crystal with catalyst like zinc, at 800° C. The grown CNT are covered with 50nm of gold using sputtering, which are masked using a special tape that loses its adhesive characteristics exactly at 120° C. The tape is used to exfoliate the CNTs from the quartz crystal and is carefully placed on the desired surface (Si/SiO₂). The wafer is then baked in an oven at 120° C, where the tape loses its adhesivity and the surface is left with gold covered CNTs. After CNTs mechanical transfer, the gold is etched from CNTs and are successfully implanted on silicon. The normal fabrication process of lithography and etching is used afterwards for making the pattern on the silicon. The process though gives an aligned solution for VLSI compatible applications, but might not be a reliable answer to the placement due to the process variations and the type of CNTs used. The mechanical transfer and CVD growth can create defects in the CNTs, changing the structural integrity of the material [25].

**3.2 Schottky Barrier Ballistic CNFET**

As current in CNFETs is formed by tunnelling of charge through the intersection of Schottky barrier created between the CNT and metal. The Schottky barrier height plays an essential role in determining the tunnelling of carriers which is dependent on the work function of the metal and CNT fermi level. Since the CNT with 1.4nm (CNTs used in this work) diameter has a fermi level energy close to 4.5eV, we have chosen Palladium (Pd).
It has a work function of 5.2eV that is in between the bands of the CNTs nevertheless it is high enough to form thinner barrier for easy tunnelling. Gold (Au) and Pd have similar work function. But, the metal carbide bonding between the CNT and Au is not as good as compare to that of CNT and Pd [30] which results in dispersion at the contact region and comparatively thicker barrier. The difference can be seen in Figure 14, hence Pd was chosen over Au.

In this work, we have not discussed the ohmic contact between the CNT and metal interface because it’s more stringent to form ohmic contact than a Schottky contact CNFET. It requires controlled environment to create a barrier, so that the metal can have work function completely aligned to the valence band of the CNTs [31]. The fabricated CNFET is best to be categorised as Schottky contact transistor even though the sample is annealed to achieve ohmic contact. Therefore, building on that and from section 2.2.3, we now extend the fermi potential of the CNT to the surface potential of the CNT-metal interface. Figure 14 and figure 15 shows better understanding of the Schottky barrier in CNFET formed between metal-carbide intersections. It can be noticed that there is no saturation or bending of the IV curve because there is no electrostatic field to interact with the sub bands of the CNT-metal surface to control the tunnelling of the charge carrier, so the device shows the Schottky contact behaviour [32].
**Operation**

**Figure 15**: Band diagram of n-CNFET with Schottky barrier, where $\Theta$ represents the Schottky barrier height for the holes to transfer from the valence band of the source to the conduction band of the drain (a) gate voltage is zero; (b) gate voltage at linear regime ($V_g > V_t$), (c) saturation region $V_D = V_g - V_t$. Adapted from [2]. Here $E_f$ is the fermi level, $V_D$ is the drain voltage, $V_g$ is the gate voltage and $V_t$ is the threshold voltage.

The operation of CNFET is similar to that of metal oxide semiconductor field effect transistor (MOSFET). Both MOSFET and CNFET have three stages of operation: cut off, linear and saturation phase. Even though CNFET and MOSFET have similar operation and structure, but MOSFETs are unipolar, whereas CNFETs are ambipolar meaning it can work as a p-type transistor and as a n-type transistor [33]. However, it is easier to fabricate p-type in the ambient condition and difficult to fabricate n-type due to the electron trapping in the air. But, the saturation operation of n-type transistor can be improved by using a low work function metal [34] like scandium for the contact metal, so that the fermi level of the CNT is closer to that of the metal’s conduction band. Thus, making the fermi level pinning easier which creates a smaller energy barrier for the electrons transition.

When no voltage is applied over the CNTs, the Schottky barrier is wide enough for the electrons that they cannot penetrate through it. And so, the n-CNFET is in OFF state. Figure 15 (a)), however shows electrons tunnelling through the drain into CNT to represents the ambipolar behaviour of CNTs. But when a small positive gate to source voltage is applied, CNFET go in the linear regime (Figure 15 (b)), where the net current is equal to the tunnelling electron difference from the source and some from the drain. As the positive bias voltage is applied across the gate, it starts shifting the fermi level of the CNTs by applying electric field across the first subband of the CNT, allowing the electrons in the CNTs to tunnel through source’s valence band to CNT’s conduction band and then CNT’s conduction band to drain’s...
conduction band; and the voltage at which the electron breaks the barrier is known as the threshold voltage of the CNFET. Subsequent to the threshold voltage, the thickness of the source Schottky barrier decreases exponentially with increasing gate to source voltage entering the third phase known as the saturation.

Difference between the electron tunnelling from source to that from drain leads to the net current of the transistor. According to Landauer model, the electrical current in a ballistic 1-D conductor is given by,

\[ I = \int_{-\infty}^{\infty} D(E) v_g(E) [f_R(E) - f_L(E)] \, dE \]  

where \( v_g \) is the group velocity, \( f_R \) and \( f_L \) are the quasi fermi functions for the right and left moving electron populations and \( D(E) \) is the density of states. Using the above equation 15, the electron drain current in a Schottky barrier carbon nanotube can be expressed as [36]:

\[ I_e = \frac{4e}{h} \left[ \int_{E_{cb} - e\phi_s}^{\infty} \{ T_{sd}(E)F(E, F_F) - T_{ds}(E)F(E, F_F - eV_D) \} \, dE \right] \]

where \( T_{SD} \) is the transmission probability from source to drain and \( T_{DS} \) is the transmission probability from drain to source. It means wider the Schottky barrier, lesser the probability to tunnel through the barrier. Therefore, in this work we have used Palladium which has a very high work function that will form thinner barrier at a low gate voltage due to fermi level pinning. Additionally, Pd has a good metal-carbide bonding [30].

During the linear regime, the current is due to source/drain tunnelling, source tunnelling and thermionic emission which is given by:

\[ I_e = \frac{4e}{h} \left[ \int_{E_{cb} - e\phi_s}^{E_{cb} - eV_D} T_{eff}(F_s - F_d) \, dE + \int_{E_{cb} - eV_D}^{E_{cb}} T_s(F_s - F_d) \, dE + \int_{E_{cb}}^{\infty} (F_s - F_d) \, dE \right] \]

\[ equation \ 17 \]
where the first term represents the transmission of electron when both the barriers from source and drain are present, the second term represent electron transmission when only one source barrier is present, and the third term represent the thermionic emission.

Neglecting the source-to-drain and band-to-band tunnelling for ideal equation, the source (drain) component for the Thermionic emission defines the $I_D$ current for the saturation region:

$$I_{TE}^{s(d)} = \frac{4q}{h} k_b T \ln \left[ 1 + e^{\frac{-E_{TE}^{s(d)}}{k_b T}} \right] \quad \text{equation 18}$$

where $I_{TE}^{s(d)}$ is the DC saturation drain current, $k_b$ is the Planck’s constant, $T$ is temperature in Kelvin. $E_{TE}^{s(d)}$ is the minimum energy with reference to the source and drain fermi level, required from the source into the channel and is given by:

$$E_{TE}^{s(d)} = q\phi_s - \frac{E_g}{2} \quad \text{equation 19}$$

where, $\phi_s$ is from equation 13, $E_g$ is the band gap of the CNTs.

Band diagram of p-CNFET and n-CNFET are complimentary to each other [35].

### 3.3 Process Variation

Like conventional transistors the CNFET are also prone to process variations, these variations can occur from wafer to wafer or within the wafer due to various parameters like variation of resist thickness, or variation in the electron gun vacuum which changes the electron beam dosage etc [37]. Apart from this, manufacturing of CNTs are also a stoichiometric process and produces a range of CNT characteristics within the same batch [38]. Thus, while designing a circuit we need to test its tolerance to such variations. Using the simulation we have displayed how variation in the length of the CNT, number of CNTs and chirality produces variation on the drain characteristics of the CNFET. For the simulation we are using the Stanford Verilog-A model, which is created for Schottky contact as well as ohmic contact [39].

In this work, we are going to study the effect of two parameters mainly, which are discussed briefly below, and will be considered while designing the memory.
3.3.1 Width and Length of Carbon Nanotube Transistor

CNT transistor is based on a model of a 1-D conductor in between an electric field, during which loosely bound electron is accelerated in direction of the field. However, some of the electrons move in the opposite direction that are known as scattering electrons. The scattering electrons depends on the electron mean free path \( L_m \) which is the average length an electron can travels before a scattering event can occur. In a perfect crystalline lattice, electrons can move without any scattering. However, in veracity CNTs’ lattice symmetry gets damaged by various sources such as impurities, defects and environmental condition [29].

In mesoscopic materials, the resistance does not scale continuously with inverse width. However, scaling of width introduces discrete modes that contributes in the transport, resulting in the increase of drain current which is completely in contrast to ohm’s law. The incorporation of bands in the electronic property can be shown by the Landauer formula (equation 15) which shows the transconductance, \( G = \frac{2e^2}{h} M T \), where \( M \) is the number of modes and \( T \) is the transmission probability of the conductor. I-V characteristics of CNFET can be represented by combining equation 13 and equation 18 and putting the value of quantum capacitance value, \( C_q = \frac{2q}{h \nu_f} \) [40]:

\[
I_D \approx 2v_F \frac{C_{ox} C_0}{C_{ox} + C_q} (V_g - V_{th}/t)
\]

\[ equation \ 20 \]
\[ I_D = g_C(V_g - V_{th}/t) \quad \text{equation 21} \]

where \( g_C \) is the transconductance of each CNT, and for each \( N \) uncorrelated CNT the drain characteristic equation can be written as:

\[ I_D = N g_C(V_g - V_{th}/t) \quad \text{equation 22} \]

The effect of change in width and length of the CNFET can be seen on the drain characteristics, figure 17.

3.3.2 Chirality

Figure 11 shows different band structure for almost same diameter CNTs. However the sub-band of the tubes differ, the main difference being in the Brillouin zone which is wider for the zig-zag tube (10,0). We can conclude that for the same number of subbands, there can be two zigzag and chiral tubes having the same energy range [41]. Which can be supported by equation 10 and equation 12 as the DOS and band gap of CNT is dependent more on the diameter rather than chirality.

Thus for different chirality, CNTs can have same surface potential but different drain characteristics (Figure 18, Gaussian distribution of chiral vector (19,0)). This is true in the absence of scattering effects and low biases, because when two CNTs have same chirality the two fold degeneracy in chiral tubes results in slight difference in the mobility of the carriers in the structure, changing the drain current [42].
3.4 Introduction to Dielectrophoresis

In section 3.1, we noticed that top-down approach is a simpler option, but it leads to unaligned CNTs with less defected CNTs, since they can be grown using arc discharge method. Whereas, the bottom-up had aligned CNTs but had more defected CNTs, as CVD grown CNTs would be used.

Dielectrophoresis (DEP) is one of the methods which provide a promising solution for placing and aligning the nanomaterials at the desired location. The principle of DEP is based on manipulating the nanomaterial in a liquid medium by applying a non-uniform electric field that leads to the polarization of the nanomaterial and their immersion solution. Due to the formation of the dipole on the object and non-uniformity of an applied electric field, the forces exerted does not compensate each other which makes a translational movement of the object.

DEP approach offers the advance of processing CNFET at ambient temperature. In the process, electric field that is generated from low voltages, using the electrodes separated by few nanometres. In addition, several parameters such as the duration of an alternating field, as well as frequency and amplitude can be adjusted to optimize the number of aligned CNTs.

The Dielectrophoresis force acting on the CNT length is expressed using [43]:

\[
F_{\text{DEP}} = \frac{\pi r^2 l}{6} \varepsilon_m \Re \left( f_{\text{cm}} \right) \nabla |E|^2
\]

\text{equation 23}

Where \( l \) is the length of CNT, \( r \) is the radius of CNT, \( E \) is the electric field between the electrode and \( f_{\text{cm}} \) is the Clausius-Mossoti factor, which represents the complex permittivity of the particle and medium. Therefore, this process depends on the conductivity of the medium and frequency of the voltage. According to alternating current electro-osmotic (ACEO), charge flow during DEP increases the conductivity across the CNT and solution in which it is dispersed. So for better manipulation we can decrease the frequency of the signal, as it increases the polarizability of the CNT and reduce any electrochemical chemical reaction.
Additionally during the process, force is majorly influenced by electro-thermal effect [44] and the van der Waals [45] forces.

The Electro-thermal force is the summation of force of attraction due to Coulomb force and dielectric medium (suspension liquid strength), which is given by the combination of both the forces:

$$F_{\text{Thermal}} = \left(\frac{\nabla \sigma}{\sigma} - \frac{\nabla \epsilon}{\epsilon}\right) \hat{E} \left(\frac{\epsilon \epsilon_0}{1 + (\omega \zeta)^2}\right) + 0.5 \mid E \mid^2$$  

\textit{equation 24}

Where, $\zeta$ is the relaxation time of particles and $\omega$ is the frequency of the pulse. Though Electro-thermal force is the summation of Coulomb force and dielectric force, Coulomb force dominates at the low frequency whereas dielectric force dominants at higher frequency.

The Dielectrophoresis circuit, shown in Figure 19, is based on the capacitive coupling of electrode, it helps in self-limiting single-nanotube assembly. The circuit is helpful in achieving the goal of carbon nanotube-based VLSI circuits. DEP provides advantages over other methods like spin-casting and chemical vapor deposition by providing more control over the position of a higher quality carbon nanotube. This method of alignment is also cheaper and time efficient option for fabrication.
Summary

Previous chapters give us a theoretical explanation of the electronic properties of CNFET and the challenges we can face while fabricating the devices. It can be also be concluded that:

a) Choosing CNT manufacturing is important to avoid unpredictable behaviour of the transistor. Transistor is prone to CNT alignment, percentage of m-CNTs and chirality along with other structural deformation during the process.

b) It’s important to understand the band structure of CNT so that threshold of CNTs can be estimated for making the electronics.

c) Different types of transistors can be fabricated by choosing from different metal for the electrode. But Palladium shows to better results because of thinner Schottky barrier.

d) Using simulations, we displayed the effect of various parameters of CNTs, mainly chirality and length, on the drain characteristics.

e) Introduction to DEP, and the parameters that can be varied.
Chapter 4: Experimental Process and Characterization

4.1 Fabrication Process Flow

Figure 20 illustrates the process flow used in preparing the electrodes and CNT deposition, which is explained in detail in the following sections. Before starting the process, the silicon wafer was cleaned using the buffer oxide etching (BOE) procedure to thermally grow 300 nm of the silicon dioxide (SiO$_2$) in the Carleton University (CU) facility. To remove oxide formation from the backside of the wafer, BOE was again performed.

Figure 20: Overview of the fabrication process used in this work
4.1.1 Lithography

Electrode patterning is done using the overlay method on Raith’s electron beam lithography (EBL). In which the first step was to do an exposure of big pattern followed by developing in MIBK 1:3 solution at 20° C for 3 minutes. Then, exposure of small patterns (nano-electrodes) followed by the same developing stage. This was done to speed up the patterning time.

Bi-layer resist was used to provide a good undercut layer and to make the lift-off procedure easier and cleaner. The recipe of the two resist Microchem PMMA 495 A6 and PMMA 950 A2 is mentioned in the Appendix.

4.1.2 Metal Deposition

Metal deposition is done using Angstrom Engineering Evaporator. Palladium (Pd) has a low adhesion with the oxide surface due low noble metal-oxygen binding energy [46]. Since, Pd had to be deposited on silicon oxide, a “glue-layer” is used in between silicon dioxide and palladium, to have a better adhesion between the two interlayer. For better adhesion two metals were tested: Chromium (Cr) and Titanium (Ti). After series of deposition and scotch tape test, Ti turned out to be a better option than Cr. However, due to the low conductivity of Ti, the glue layer adds to the parasitic.

4.2 Preparation of CNTs Solution

CNTs tend to form bundles in the order of tens of micro-meter, and comes in as metallic and semiconductor. Therefore, it becomes difficult to separate out individual CNTs without functionalizing them. In this work, we tried to use different solutions to disperse the CNTs with an aim to keep the CNTs pristine and make single CNT device. Based upon the best dispersion of the CNTs we have used NMP (N- methyl-2-pyrrolidone) after trying various other organic solvents like chloroform, 1,2-dichloroethane, and IPA. After series of deposition and dispersion, we can suggest that NMP is suitable for dispersing the Raymore Nanointegris’s SuperPure CNTs. The only observed disadvantage of NMP is the high vaporization temperature which makes the deposition more variable, due to the Brownian motion of the CNTs.

For preparing the solution, Ultrasonic bath (Elmasonic P- sonic bath) was chilled by using the ice pack till it reached 11° C. This was done to mitigate the exothermic nature of the process which can cause temperature of water to rise up to 40° C. Then to have maximum control on
oxygen-carbon interaction, the CNTs were only taken out in the vacuum therefore the whole process of cutting CNTs was done in a glove box.

Note: Variation in initial and final temperature of the solution leads to a different amount of dispersed CNTs due to the van der Waals forces between the CNT. Therefore, for every deposition if it took more than 1 day, the process was repeated.

4.3 Electrode Design

Generation of the CNTs is a stochastic process and involves separation of metallic CNTs (m-CNT) and semiconducting CNTs (s-CNTs). The design of the electrode is based on increasing the probability of CNT landing on the electrode.

CNFET is dependent only on the semiconducting CNTs [47], therefore the probability of the CNFET ($P_{CNFET}$) is equal to the number of CNTs ($n$),

$$P_{CNFET} = (P_{semi\_cnt})^n \quad equation \ 25$$

where, $P_{Semi\_cnt}$ is the probability of the semiconducting CNT.

Since the bundle of CNT is a mixture of metallic and semiconducting CNTs, therefore the mean of the current is given by

$$\mu_s = \int_0^\infty I_s(x)f_{sd}(x)dx \quad equation \ 26$$
$$\mu_m = \int_0^\infty I_m(x)f_{md}(x)dx \quad equation \ 27$$

where $\mu_s$ and $\mu_m$ is the mean of semiconducting and metallic current flowing from source to drain CNT, respectively; $f_{sd}$ is the fermi direc function of semiconducting CNTs and $f_{md}$ is the fermi direc function of metallic CNTs [48]. So, in a transistor if there is a mixture of metallic and semiconducting CNTs, then the mean current of the transistor can be expressed as

$$\mu(I_{CNT}) = p_s\mu(I_s) + p_m\mu(I_m) \quad equation \ 28$$

Where $p_s$ and $p_m$ is the average fraction of the semiconducting and metallic CNTs.
The design of the electrode is conceptualized on the interdigitated resistor and relays [49]. The x-direction (row) provides an increased probability of combating shorts, so it makes sure that there is at least one CNT in a series is semiconducting. Therefore, the probability of finding a semiconducting CNFET is given by:

\[ P_{\text{row}} = 1 - (1 - P_{\text{CNFET}})^x \]  
\[ \text{equation 29} \]

Where \( x \) is the number of CNFET in series. Row connection increases the probability of placing a semiconducting CNT within the electrode, however it decreases the current drive.

Consequently, it becomes important to strategically employ correlated and uncorrelated redundancies so that the CNFET is more metallic-CNT tolerant and gives high current drive [47]. Therefore, if we increase the number of column the current drive will commensurately increase. The column CNTs itself can give a high current drive but it will not be tolerant to the metallic CNTs, as a single metallic CNT can create a short between the drain and the source. Therefore, our electrode design uses asymmetric correlation to achieve metallic tolerant design.

Figure 21, shows the fabricated electrode and; illustrates how row connection is forming a series of stack transistor and column connection is forming a parallel connection. Electrode
layout provides a symmetric design, however the electrode is still susceptible to process variations during lithography, metal deposition ... Nevertheless, the nanoelectrode design is reproducible with 5% of variation and 90% yield from wafer to wafer using the parameters and PMMA resist recipe mentioned in the Appendix.

4.4 Dielectrophoresis

For finding suitable frequency, we performed finite element analysis (FEA) of the electrode design using HFSS shown in Figure 22. It was noted that as we increased the frequency of the DEP, the electric field intensity around the sample augmented, especially around the corners of the nanoelectrodes where it increased significantly. Figure 22 (a) shows DEP simulation at 50kHz, at that particular frequency stray electric field intensity is minimum and a uniform field is observed between the electrode, where the CNTs are present. Thus, helping in the CNTs’ orientation. Whereas, in Figure 22 (b) when the frequency increases from 50 kHz to 150 kHz, the electric field simulation becomes non-uniform mainly due to parasitic electric field. This parasitic electric field is high around the corners of nanoelectrode which can be seen through the simulation. When a nanoparticle is suspended in a liquid, this parasitic electric field will attract more CNTs, accordingly producing undesired results. This parasitic electric field increases, as we increase the frequency for the deposition, figure 22 (b). This parasitic/stray electric field is due to the high frequency of the signal applied which decreases the impedance at the corner of the electrode, resulting in capacitive coupling.

The DEP depends on the thickness of the oxide and distance between the electrodes, because the circuit connection was chosen with an aim to provide capacitive coupling to the adjacent electrode to provide a less intense electric field.

Figure 23, shows the experimental result of DEP performed at 50kHz and 150kHz on 300nm SiO₂, which reflects the results from the simulated FEA. Table 1 shows the parameter used to model the CNT within the HFSS. During the simulation, we didn’t consider the Brownian motion which can lead to certain unsolicited CNTs and duration of the signal, since as the time increases the dipole across the number of CNTs increases and thus a greater number of CNTs are attracted. We did iterations to find the suitable time for which the signal should be applied and kept the evaporation time of the NMP to be constant at 5 minutes, after which we transferred the sample on hot plate. The result for different period time is shown in Appendix Figure 58, Figure 59 and Figure 60. As duration of signal increased, proportionally CNTs got attracted towards the nanoelectrode.
Table 1 Parameters used for simulating CNTs in HFSS

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Value</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$L$</td>
<td>Length</td>
<td>100nm</td>
<td></td>
</tr>
<tr>
<td>$r$</td>
<td>Radius</td>
<td>1.4nm</td>
<td></td>
</tr>
<tr>
<td>$\sigma_{_{SWNT}}$</td>
<td>Conductivity of semiconducting CNT</td>
<td>2.88 S/m</td>
<td>[50]</td>
</tr>
<tr>
<td>$\sigma_{_{m\text{-}}SWNT}$</td>
<td>Conductivity of metallic CNT</td>
<td>103 S/m</td>
<td>[50]</td>
</tr>
<tr>
<td>$\varepsilon_{_{\text{semi\text{-}}CNT}}$</td>
<td>Permittivity of semiconducting CNT</td>
<td>$5 \times \varepsilon_0$</td>
<td>[51] [52]</td>
</tr>
<tr>
<td>$\varepsilon_{_{\text{metallic\text{-}}CNT}}$</td>
<td>Permittivity of metallic CNT</td>
<td>$103 \times \varepsilon_0$</td>
<td>[51]</td>
</tr>
</tbody>
</table>

Where $\varepsilon_0$ is the permittivity of the air.

![Electric Field formation around the CNTs at (a) 50kHz (b) 150kHz](image1)

*Figure 22 Electric Field formation around the CNTs at (a) 50kHz (b) 150kHz*

![SEM image of the Transistor using the DEP (a) at 50Hz, 2V_{p-p} and 5 secs (b) at 150kHz, 5V_{p-p} and 5 Secs](image2)

*Figure 23 SEM image of the Transistor using the DEP (a) at 50Hz, 2V_{p-p} and 5 secs (b) at 150kHz, 5V_{p-p} and 5 Secs*
4.5 Annealing

Annealing process acted as an important step in enhancing the performance of the CNT transistor, but determining the optimum temperature was a task. If the temperature is above 800° C the metal for the contact becomes too soft and becomes useless or seep into the silicon. Whereas, if it’s low the metal-carbide bond is not formed and the contact resistance between CNT and metal is not reduced because of wetting effect [53].

Annealing for this process is done using the Rapid Thermal process at 350° C for 5 minutes in the presence of nitrogen and argon. The process for the annealing is based on a quick ramp of temperature; a steady flow of N₂ and Ar in a ratio of 1:1 for 5 minutes; and a rapid cool down.

4.6 Characterization

![Image](image.png)

*Figure 24: (a) Pictorial Representation of the characterization setup (b) Device under test (DUT) snapshot*

The characterization setup, shown in Figure 24, is made using the Everbeing Probe station having titanium probe with 5 um which is connected to a system for monitoring the position of the probes. Keysight B2900A was used for measuring and a constant DC supply (Korad 3005D) is used for supplying the gate voltage. The designed electrode displays good results and exhibited a good Schottky and ohmic contact transistor with reproducibility. Figure 25, shows the I₀-Vₐ characteristic of the transistor and displays good ohmic region with low output current before annealing, but after annealing the output current is increased significantly.

Even though the IV curve shown in Figure 25 does not perfectly match with the IV characteristics shown in Figure 17, the device characteristics are similar to that of MOSFET device. It showcases the triode regime, i.e., it consists of three regions: Cut off, Linear (Triode) and saturation. These regions are dependent on the gate-source voltage and drain-source voltage. And the main reason that the transistor didn’t reach the saturation, is because the sample consists of only 33% of semiconducting and rest metallic CNTs. And therefore,
increase in the drain voltage leads to the heating of the CNTs which burns off between 4.5-5 V. This is also known as the electrical annealing. It is a good method for removing the metallic CNTs (m-CNTs) from the semiconducting one [54], but it also damages the semiconducting CNTs.

Figure 25: (a) Drain characteristics with the gate voltage (5, 10, and 15V) showing the n-type CNFET, before annealing (b) Shows the drain characteristics after annealing at 350° C.

Furthermore, we can notice that the p-CNFETs (Figure 26) shows better saturation region than n-CNFETs, this is because of the oxygen charge trapping which under the ambient atmosphere makes electron deficiency in the CNTs [53].
Figure 26: (a) Drain characteristics with the gate voltage (5, 10, and 15V) showing the p-type CNFET, before annealing (b) Shows the drain characteristics after annealing at 350° C.

Figure 27: Simulation and experimental drain characteristics of the fabricated CNFET.
Linear Region

The fabricated transistor is capable only to work above the threshold voltage of the CNFET, that is, where it enters the linear region. At the linear region, the current flowing through the Schottky Barrier is not mainly affected at the source end while varying the gate voltage. Instead, the tunnelling in the barrier is controlled by the drain side of the transistor. Therefore as the source to drain voltage increases, the drain Schottky Barrier decreases in width and there is a linear increase in the current. As shown in Figure 25 (a), we can notice that as we increase the gate voltage, the slope of the $I_d$ v/s $V_{ds}$ curves increases linearly in the linear region. The slope is proportional to the rate of change of current over the drain to source voltage.

Saturation Region

The point when no barrier exists at the drain end of a CNFET the saturation region starts. Figure 26 (a), shows that saturation current for 10V gate voltage is 4nA after which the current remains relatively constant. The tunnelling at the source becomes constant and therefore there is no proportionate change in the current.

Due to the thickness of the oxide, the designed electrode will not reach the saturation. Since the oxide thickness is very thick (300nm) the gate voltage does not have a large control over the IV characteristics as compared to a thinner gate oxide [54]. Additionally, the drain to source voltage needs to be highest in order to have large output swing however in that case the CNTs burns off.

Using the Stanford model and replacing the values from the measured transistor (data mentioned in the Appendix): refractive index of the Oxide ($\varepsilon = 2.22$), oxide thickness=300nm, top gate refractive index = 1 (Air) and substrate capacitance = 3.442pF, the comparison of the measured and simulated value is shown in Figure 27. However, the IV characteristics shown in Figure 27 doesn’t shows good saturation region, this is because the sample was a mixture of m-CNTs and s-CNTs in a ratio of 1:49, and that is why the electrode was formed in an interdigitated configuration. Also, the D peak from the Raman data shows the structural damage of used CNTs.
Conclusion

The thesis presented a process flow for making a back gate CNFET which can be innovative for fabricating VLSI application based on CNTs. The presented CNFET is fabricated using electron-beam lithography on a bi-layer resist followed by metal deposition and characterized using two probe method, SEM and AFM. Subsequently, CNTs are deposited on the electrode using the DEP using low-frequency signal, and the finite element analysis in ANSYS HFSS showed how electric field is formed around the CNTs which helps in aligning the carbon nanotube. The fabricated CNFET electrode had a reproducibility of 90% (wafer to wafer). Moreover, the interdigitated electrodes design help in making metallic CNTs tolerant CNFET while having a good current drive.

DEP showed a prospective result for fabricating beyond CMOS transistors since the process is based upon nanomaterial manipulation. The result showed different effect depending on the parameter like frequency of the applied signal and the concentration of the solution with a

Figure 28: Flowchart summary for the DEP process
deviation of ± 15%. 11 out of 15 devices fabricated using DEP process showed a promising bid.

The significant change in the IV characteristics was observed after the annealing process at 350° C, as it reduced the wetting of the CNTs. Additionally, this step was also beneficial for increasing the metal-carbide bonding and decreasing the contact resistance, following which the field-effect behaviour of the CNFET showed the better current drive. Therefore, to decrease the contact resistance annealing was an essential step.

Further advancing the process mentioned in figure 28, it can produce VLSI based circuit or a sensor at a lower cost. Because, no dedicated CVD system is required for CNTs.
Chapter 5: Application as a Memory: Static Random Access Memory

5.1 Introduction

CMOS technology till now has been delivering densely integrated, low power and fast response circuits [55]. However, to accomplish that goal additional circuit overhead and circuit reliability has become major concerns. Static Random Access Memory (SRAM) which occupy approximately 84% of the total area in a System on Chip (SoC) has a major effect on the overall performance of the SoC [56]. It is also the fastest and comparatively expensive memory used in the computer architecture. Therefore, to maintain smaller area and power consumption in the nano-regime while commensurately maintaining the stability to the noise and process variation extensive efforts are made for improving the design of the SRAM [57]. The problem intensifies with the upsurge of upcoming Internet of Things (IoT) which are being programmed to run data intensive programs like a pre-trained neural networks based on Von-Neumann computer architecture. The von-Neumann architecture is based on two way communication between the embedded memory and the processor. This inter-communication creates power and throughput inefficiency. To alleviate this problem, we had earlier proposed an unconventional Von-Neumann architecture based on Computational Random Access Memory (CRAM) [58] using MOS technology. It could perform logical AND and OR logical function as well as conventional memory operation. This eccentric solution memory architecture reduces the data communication between the memory and the processor. These novel approaches aid in the integration of the memory and processor [59], thus supporting faster processing for application like machine learning [60].

Section 5.2 gives the introduction for SRAM and explains the working of the conventional SRAM based on MOS technology. Subsequently, section 5.3 presents a small introduction to CRAM technique used in the past using MOS technology. Proceeding from MOS to CNFET; section 5.4 discusses the proposed SRAM using CNFET, including its layout based on DEP. Furthermore, the proposed SRAM is designed to implement in-processing logic circuits for AND, OR and XOR.
5.2 Static Random Access Memory

While maintaining small area and reliability, scaling of SRAM is done to reduce the cost per bit and fast operating cycle. Efficient SoC memory cells typically consists of a positive feedback cross-coupled invert latch which can perform read and write operation. The positive feedback loop is controlled using the access transistor, which are connected to the bit-lines. These bit lines are the input/output lines for the memory cell, as the cell charges or discharges from the cell using bit-lines. For example, when the access transistor are ON, these bit-lines are made complementary to each other so that the differential voltage can be used for activating the positive feedback loop and the desired voltage can be saved in the cross-coupled inverter. Besides SRAM bit cell, the memory architecture consists of additional peripheral circuits, such as address decoders, which steer desired operation into the required memory address, bit-line pre-charge circuit, sense amplifiers and input/output driver.

*Six transistor SRAM bit cell*

As mentioned earlier, SRAM bit cell consists of cross-coupled inverters. In the conventional architecture six transistor (6T) SRAM bit cell (Figure 29) P1-N1 and P2-N2 forms the cross coupled inverter, whereas N3 and N4 forms the access transistor for the SRAM bit cell. When the access transistor is ON the charging and discharging of SRAM bit cell takes place through nodes A and B. To maintain the positive feedback and to avoid flipping of the cell, certain sizing rule is followed. For more data stability, both the pull down devices (N1 and N2) from the inverter should be stronger than the access transistor. Whereas for the better write, the pull up transistor (P1 and P2) should be stronger than the access transistors.

![Conventional 6T SRAM bit cell architecture](image.png)
The access transistor is controlled using the wordline (WL) signal attached to the gate terminal, whereas the drain terminal of the transistor is connected to the bit cell, the source terminal is tied to the bit line (BL or BLB) [61].

For read operation, the access transistors are turned ON, whereas the bit-lines (BL and BLB) are charged to VDD. As a result, there is a drop in the voltage within one of the bit-lines creating differential voltage that is read using the sense amplifier. For example, if node A is storing bit 1 and node B is storing bit 0 then bit line BLB will start discharging via node B using N2. The feedback loop is maintained using positive feedback. Whereas, for write operation, access transistors are turned ON but bit-lines are made complementary to each other. Due to this configuration of the bit-lines, the positive feedback gets weaker and value stored in the cell is toggled.

**Eight Transistor SRAM bit cell**

One of the disadvantage of 6T SRAM bit cell is the common access transistor for, both read and write operation. Due to this, 6T SRAM had lower noise margin for read operation, making it more prone to the noise. Therefore, to circumvent this problem dual port SRAM bit cell is proposed using Eight transistor (8T). Here, two extra transistors (N5 and N6) are used for creating the read port for SRAM bit cell. Transistor N6 is controlled using the read worldline (RWL), its source terminal connected to a bit line (RBL) dedicated only for read operation and its drain connected to another transistor which provides a discharging path during the operation. For example, for reading stored bit 0, RBL is given high signal which turns ON the transistor N6 (Figure 30), since the node A is storing 0 and node B is storing 1. Therefore, node B makes the

![Figure 30 Conventional 6T SRAM bit cell architecture](image)
transistor N5 work in the saturation phase, making a path for the RBL to discharge. However, during read 1, node B is storing bit 0 which keeps transistor N5 in the cut-off region. Therefore, it does not allow the RBL to discharge and keeps its signal high.

\textit{Computation RAM}

Computational Random Access Memory (CRAM) provides an analog solution near the memory to perform the logical operation. Several publication had been published with different names like in-computing memory and processing in-memory [62][63][64]. The main objective while designing these memories is to reduce the latency, process variation and additional overhead, so that maximum usage of memory’s bandwidth can be utilized [65].

These modifications of the conventional memory provide a simple interconnection between operational processing elements and the memory, which will increase the computational capability inside the chip without any major modification. However, some challenges may arise for the programmers. For example new data structure for high performance CRAM which is better than parallel data structure on multi-core machines [66]; rigorous analysis methodologies to provide easy and realistic simulation using existing infrastructure; creating coherence between the processing unit and CRAM for the shared data [67].

\textbf{5.3 CNFET based Single Ended SRAM with Processing-in-Memory}

Though the earlier proposed CMOS SRAM is capable of fulfilling the data sensitive AI application, it isn’t capable of performing the XOR operation which is essential for performing computation in neural networks [68]. Also, the single-ended CMOS SRAM cell in MOS technology has a higher read access time [69]. Therefore, we propose a new bit cell which is based on Schmitt Trigger operation. Using the CNFET Schmitt trigger we can mitigate the disadvantage of higher read time and comparatively faster write operation which we faced during the CMOS designing. Figure 31 shows the Voltage Transfer Characteristics (VTC) of CMOS inverter, CNFET inverter, Schmitt trigger CMOS inverter and Schmitt trigger CNFET inverter, showing the fast switching from high to low state using the beyond CMOS material and a faster discharging using the CNFET Schmitt Trigger inverter. This simulation is done using the bulk CMOS 32nm PTM model file (http://www.eas.asu.edu/~ptm) and 32nm CNT channel length.
While designing an SRAM bit cell, certain design constraint needs to be followed so that while charging and discharging of the storage node, it does not lose its data integrity. In the case of MOS technology, transistor’s width to length ratio is a primary criterion [70]. However, it’s not directly applicable to the SRAM using CNFET since there is no direct 3-D channel. So analogous to the width of the MOS transistor CNFET-SRAM can be designed using different number of other parameters which have the same impression of controlling the charging and discharging of the current. The first parameter is the CNTs per transistor which changes the width of the transistor.

![Figure 31: Voltage Transfer Characteristic (VTC) of CMOS and CNT based inverter and Schmitt Trigger Inverter](image)

While designing an SRAM bit cell, certain design constraint needs to be followed so that while charging and discharging of the storage node, it does not lose its data integrity. In the case of MOS technology, transistor’s width to length ratio is a primary criterion [70]. However, it’s not directly applicable to the SRAM using CNFET since there is no direct 3-D channel. So analogous to the width of the MOS transistor CNFET-SRAM can be designed using different number of other parameters which have the same impression of controlling the charging and discharging of the current. The first parameter is the CNTs per transistor which changes the width of the transistor.

![Figure 32: CNFET model used for simulation which is based on top gate formation](image)
This parameter also governs the total area in the CNFET which is expressed in terms of physical gate width \( W_g \), where \( W_g \) is given by [48]:

\[
W_g = s \times (N-1) + d_{CN} + 2 \times W_{OV}
\]

\[\text{equation 30}\]

here, \( s \) is the pitch between the CNTs, \( N \) is the number of tubes in a CNFET, \( d_{CN} \) is the uniform nanotube diameter, and \( W_{OV} \) is the overhang width of the gate from the edge. Secondly, the strength of a CNFET \( (I_{ON}/I_{OFF}) \) can also be adjusted by using multiple threshold CNTs [72], which can be achieved by using different chirality CNTs. Since chirality governs the threshold of the CNFET by changing the diameter and type of the CNTs,

\[
V_t = \frac{\sqrt{3} a V_\pi}{3e D_{CNT}}
\]

\[\text{equation 31}\]

where, \( a = 2.49\text{Å} \) which is the carbon-to-carbon atom distance, \( V_\pi = 3.033 \text{ eV} \) is the carbon \( \pi-\pi \) bond energy in the tight binding model, \( e \) is the unit electron charge, and \( D_{CNT} \) is the CNT’s diameter. The disadvantage of using multiple chiral CNFET for designing a circuit is increasing the process variation and area of the circuit. Because growth of CNTs is a stochastic process so getting a precise CNTs is difficult and placing those different chiral CNTs at specific location is also challenging. This can be made easier only by allocating separate area for same chiral CNTs which in turn increases the area of the circuit.

This work for designing the SRAM is based on simulation using the Stanford model, which is shown in Figure 32. The model is based on CVD grown CNT and top gate fabrication process with dual gates.

Figure 33: Hysteresis loop formed by the Schmitt trigger used in the proposed SRAM bit cell
While performing the low voltage circuit operation in scaled-down technology, the circuit design has a severe effect from the process variations which increases the probability of the failed read/write operation. And since the SRAM covers a major portion on the system on-chip (SOC), designing cell which are less prone to the errors due to process variations and noise becomes essential. Therefore, we implemented the use of single-ended SRAM using the Schmitt trigger-based inverter which shows less variation to the CNTs’ characteristics like chirality and can work on multiple voltages with a good noise margin. Schmitt trigger based inverter is a bi-stable device, which has two different switching voltages. Its switching voltage is dependent on the input and output which is produced by its hysteresis, shown in Figure 33. This hysteresis forms the basis of the noise-immunity characteristics of the Schmitt trigger [73].

The final scope of this work is to implement SRAM bit cell based on non von Neumann architecture based memory architecture. The 6T bit cell which is a widely used SRAM bit cell architecture (Figure 34 (a)) is not preferred for designing because of its requirement for differential signal in reading and writing operation which adds an additional overhead for producing differential I/O. The conventional 6T CNFET architecture consisting of two inverters forming a bi-stable circuit similar to that in a MOS technology that is highly accepted in the industry. The circuitry is built upon the mutual charging and discharging of two storage node (X and XBAR) to perform three operations Read, Write and Hold. During write operation, two bit lines (BL and BLB) are given differential signal to activate positive feedback within the circuit that helps in maintaining the storage. For example during write 0 operation, assuming previously stored value is 1, the BL is pulled down to 0 whereas the BLB is pulled up to 1 and subsequently the access transistors (WWL) are turned on which allows the currents to discharge from voltage node X to the ground, while from the complementary side which, is the XBAR node, is charged, maintaining the positive feedback that helps in recovering from the voltage drop at the node X.

For reading operation, the BL and BLB are pulled up to 1 and the WWL is turned on which creates a voltage drop in one of the bit line which is subsequently sensed through a sense amplifier. This is achieved by discharging one of the bit line, through the node voltage where 0 is stored, that forms a path for the corresponding bit line to discharge into the cell, whereas complementary port provides the feedback enough to withstand the voltage spike.
However, since the discharge is done within the circuit it makes the bit cell more susceptible to the noise. To make feedback loop invulnerable to such noise, large pull-down transistor is required for designing the inverter so that the discharging is faster than charging of the storage node, which if failed can lead to the flipping of the memory stored.

Though the 6T provided a stable solution for write operation, it lacked the immunity from noise while performing a read operation, therefore to resolve the problem the second most used architecture in the industry is 8T SRAM bit cell, shown in Figure 34 (b), which has a dual port for the read and for the write. While the write operation remains the same, the read operation is performed by disintegrated transistors (N7 and N8) which are controlled via Read Word line (RWL) and Read Bit line (RBL) signals.

During the read operation, RWL is turned ON and RBL is pulled up to 1, now depending on the storage node voltage the N7 will be ON or OFF providing a path for the RBL voltage to discharge out of the bi-stable positive feedback circuit cell. For the read operation the 8T is dependent on single bit line whereas for the write operation the cell requires two bit line, however, the 6T bit cell requires two bit line for the read as well as for the write operation which provides a disadvantage of slower operation between the I/O.

The single ended operation proves to be a beneficial attribute in the SRAM memory because using this uncorrelated mechanism, the data can be read from one port and the other port can
be used for writing while reducing the circuit overhead trace from the layout, but this type of solution has not been widely used in the MOS industry due to high power consumption but with rising of new material like CNFETs which have lower power consumption it can be a very useful technique for memory design and will provide a novel architecture solution to the computing. Because using the one port of the single-ended memory data can be directly computed, and the other end can be used for writing the data faster without being converted to differential signals, thus reducing the trace length and resulting in faster computation. This type of computing is the most promising approach for a unique architecture known as in-processing memory where the logical computation is performed within the embedded memory in order to reduce the memory-processor data transfer and utilizing the full bandwidth of the memory, thus providing a conceptual solution to the bottleneck problem, shown in Figure 35. The advantages are not limited only to the throughput efficiency but also has a power efficiency factor included to it, due to the mentioned advantages it has a significant impact on the computer architecture. Various architecture has been proposed in the past using the MOS technology. D.G. Elliot et al proposed integrated logical unit close to the memory to reduce data latency [74] and most recently Amogh et al [75] presented a similar solution of integrating the logical unit inside the memory to increase the utilization of the memory’s internal bandwidth. But in doing so their proposed architecture raises the ground voltage of the chip, which will make the cell RSNM poor and more susceptible to noise, especially when the read operation is used for the logical operation. Moreover, it poses a greater risk to the half-selected bit cells. The process in-memory concept is not confined to conventional MOS technology but also being investigated as possible architecture using beyond CMOS material like phase change material. The aim for such architectures is to meet the need for data intensive AI applications [76]. Therefore with ballistic electronic transport and low power consumption CNFET can be one of the popular beyond CMOS technology to be implemented for such architectures.

The following work has been organized as follows, Section 5.2.2 discusses the proposed SRAM cell architecture, Section 5.2.3 discusses the characteristics of the proposed SRAM bit cell along with the previously published CNFET based 6T and 8T SRAM cell, Section 5.2.4 shows the layout for the proposed SRAM bit cell and the procedure which can be used for fabricating it and finally Section 5.2.5 covers the proposed SRAM bit cell as processing in-memory architecture.
5.3.1 Proposed SRAM bit cell

With the MOS technology the single-ended SRAM has been neglected because it created a write latency problem due to the absence of complementary Bit line (BL) signals. The single-ended voltage charging/discharging took time to flip the stored value. But with new materials like CNFET which have no phonon dispersion and higher near-ballistic properties [77], the single-ended SRAM can be implemented and will be very useful in developing faster memory. Since single ended SRAM only requires one BL for writing and one BL for reading (RBL), it reduces the peripheral circuit trace for generating the differential I/O which can be omitted. Additionally, this decoupling of the read and write ports proves to be an improvement in the Read-SNM for the proposed SRAM bit cell. Since the CNFET characteristics are sensitive to the chirality of the CNT, the proposed SRAM shown in Figure 36, is designed using the Schmitt trigger which makes the circuit stable for multiple low voltages and immune to process variations due to the built-in hysteresis. It’s further enhanced by a dynamic feedback mechanism. The feedback in the Schmitt trigger is designed in such a way that the write operation can be easily achieved, which is usually an onerous operation in single-ended SRAM, without compromising the area.

![Proposed SRAM bit cell](image)

*Figure 36: Proposed SRAM bit cell*

The Schmitt trigger decreases the slope of the Voltage transfer Characteristics (VTC) which results in the lower trip voltage of the inverter. Furthermore, to weaken the pCNFET during the write operation the Schmitt trigger feedback loop transistor (N6) is sourced with the word line so that during the write 0 operation when it is ON, the feedback nCNFET (N6) is sourced with VDD. As a result during the write operation, it raises the voltage between N4 and N5 by VDD-V_T. Whereas, during the read 0 operation when the current from the bit line discharges
into the cell, and the N6 acts as a voltage dividing element by reducing the read current entering into the cell, thus has good RSNM. The proposed SRAM bit cell is made using a single chiral CNT with chirality (19,0), diameter approximately 1.5nm. The number of CNTs per transistor for the SRAM is N1(2), N2(1), N3(2), N4(1), N5(1), N6(1), P1(1) and P(1).

**Write**

For write 0, the BL is pulled down during the operation and subsequently the P2 transistor is switched ON, which makes the coupled inverter’s N2 transistor and the Schmitt trigger’s N6 transistor change from OFF state to ON state. Since N6 is connected in between the N4 and N5, which are transiting from ON to OFF; current does not go through N5 easily as N6 drain air resistance by supplying VDD-Vt between N4 and N5. This creates a slow discharging at XBAR which in turn changes the state of N2 faster by creating a faster threshold electro-potential which affects the subbands of the nCNFET.

![Figure 37: Write operation with varying BL voltage to show the respective transient signal integrity](image)

Thus, by creating a virtual ground for N4, the circuit gets easily discharged [78]. In other words, we can assume that the larger current flows through P2 creating a voltage hike and making N2 to flip from OFF to ON that changes the state from X to 0. This large current flow also has a disadvantage because it increases the power consumption during write 0 operation. Similarly, in write 1 operation, transistor N4 and N5 are turned ON as the BL is pulled up to
1V. As the state of Xbar transient from 1 to 0, P1 start reaching the threshold of the coupled inverter which then starts transiting from 0 to 1, hence charging node X.

Figure 37 shows the signal integrity of the node voltage with different BL voltages. While the WWL remains the same, it can be seen that the tripping point for the proposed bi-stable circuit is close to 330mV. That means to achieve a full charged node (1) the minimum voltage required is more than 330mV, below that the circuit doesn’t activate the positive feedback loop resulting in data loss. This analysis also demonstrated the dynamic write ability and the advantage of using the Schmitt trigger in the SRAM cell as now the memory cell can operate under different voltage under limited pulse width and amplitude. The dynamic write ability can also be associate it with the real-time operation, and it becomes an important parameter when working under high-frequency clock.

Read

The proposed SRAM performs reading operation using the transistor N2 and signal RWL. When the proposed cell is operated for read function, the read bit line (RBL) signal is pre-charged to VDD, RWL is transient from 0 to 1 which switches ON the N2 transistor while WWL is disabled, assuming a read 0 operation, the decoupled read port turns ON the N6 and the RBL discharges into the ground. However, during the read 1 operation, transistors P1, N4, and N5 are turned ON, subsequent to the RWL signal N2 is turned ON and Schmitt trigger’s positive dynamic feedback comes into action. As the RBL start discharging via N2, N6 starts attaining its threshold voltage and the read port forms a voltage divider circuit, helping in maintaining the data integrity.

The disabled WL and the decoupled BL made the RSNM of the memory bit cell equivalent to the hold SNM.

Hold

The hold function for the proposed SRAM bit cell is performed by disabling all the signals. During hold 0, N6 switches ON thus making the read port comparatively stronger than the counterpart inverter’s n-CNFET. Thus to mitigate the asymmetric design the number of CNTs used in N3 is equal to the sum of CNTs used in N4 and N5. While the hold 1 operation is unpretentious as N6 is switched OFF, the charge is stored in the bi-stable inverter circuit in which read port has a stacked n-CNFET load.
5.3.2 CNFET SRAM Bit Cell Performance

Verilog-A CNFET parameter modelled by Stanford University's Nanoelectronics Group [79] [39] has been used for simulation. The simulation is accurate to the fabricated CNFET till 1.1V. Since the values used for fitting are small it leads to convergence error due to large computation. Therefore, for this work we have used original model with supply voltage at 1V. The nominal chirality in the simulation is (19,0) which gives a close diameter of what we have, and the pitch of the CNTs is 20nm. The number of CNTs per transistor for the conventional 6T SRAM are N1(2) N2(2) N3(3) N4(3) P1(2) and P2(2), whereas for the 8T conventional cell the configuration is the same with 2 additional CNFET N5(1) and N6(1)[80].

Write Ability

In the single-ended SRAM, the write operation is done by charging and discharging of the circuit using only one voltage node which makes it difficult for performing the write operation. In conventional 6T and 8T SRAM, the write operation is performed using two voltage storage nodes having differential signals. Therefore, for finding the write stability and immunity to noise the analysis would depend on both the ports.

Conventionally, write stability would be measured by writing on a single port (displaying its charging operation) and reading it through the second port (displaying its discharging operation). However, it is not the case when designing the single-ended SRAM as this correlated charging and discharging is not applicable. Rather it’s an independent operation for both the ports reliant on the write trip voltage of the inverter. Therefore, the conventional write static noise margin (WSNM) [81] wouldn’t be a practicable methodology to characterize the writing ability. Various other methods have been implemented in the MOS based single ended SRAM circuits like the bit line sweep method [82] and the word line sweep method [83]. The electro-potential across the CNT plays an important role since the Schottky barrier is present at the CNT and metal interface, which can be controlled by applying the electric field across the CNT that shifts the first subband of the CNT making a thinner barrier to cross. Therefore, using the word line sweep method will be more valid to the CNFET. It will have realistic regulation over the threshold voltage of the CNFET by affecting the electrostatic tunnelling [32]. The word line sweep method is performed by applying word operation condition (assuming the data stored within the cell is 1). That is, WBL=0, RWL=0 and RBL=0 and then sweeping the WBL from 0 to 1.
During this transition, the CNTs’ electronic bands start degenerating and reaches the two quanta of the conductance, and toggles the X and XBAR at the bit cell trip voltage. WSNM is given by the difference between the supply voltage of the SRAM bit cell and the trip voltage. Figure 38 (a) shows the analysis configuration and Figure 38 (b) shows the comparison between the conventional 6T and 8T SRAM CNFET and the proposed SRAM bit cell. It also verifies the trip voltage (330mV) of the circuit mentioned before. The proposed SRAM uses the Schmitt Trigger for forming the hysteresis that helps in having a low trip voltage, increasing the difference between the supply voltage and the trip voltage. Thus, making it more immune to the static noise that can be generated during the write operation. Additionally, since the writing mechanism is identical for the 6T and 8T they display the same WSNM.

**Read Ability**

Read operation is performed by discharging the read current from the RBL, which increases the probability of flipping the stored value in the memory. Therefore, to measure the immunity of the SRAM cell during the read operation, butterfly curve is simulated by plotting the VTCs
of the inverters, Figure 39 shows the butterfly curve, that represents the Read SNM (RSNM) of the memory cell and is calculated by fitting the largest square inside the lobes of the curve.

![Butterfly Curve](image)

*Figure 39: Comparison of Read SNM between proposed SRAM bit cell and conventional SRAM bit cells*

For simulating the RSNM, the read configuration is applied on the cell, i.e., WWL = 0, WBL = 0, RBL=1, and RWL=1. However, since the conditions are static and the curve is formed from the VTCs of the inverters, it is achieved by applying noise at the storage node which transients from 0 to 1. The analysis setup for performing the Read Static Noise Margin is shown in Figure 40.

![Read SNM Simulation Configuration](image)

*Figure 40: Read SNM simulation configuration*

The proposed memory cell show 60% more RSNM than the conventional 6T due to the disintegrated read port. However, this design doesn’t allow all of the current to bypass the bistable circuit which decreases the RSNM by 10% compared to the conventional 8T SRAM cell.
Moreover, the proposed SRAM bit cell design uses the Schmitt trigger which provides charging and discharging through hysteresis. Resulting in less read current. Figure 41, shows the comparison of performance between the conventional SRAMs and the proposed SRAM. We can notice by applying the Schmitt Trigger and dynamic signalling there is substantial drop of 73% in the read current and 4% in the dynamic power during the write 1 operation. This can be attributed to high ON/OFF current ratio of CNFET [84].

**Figure 41: Performance comparison of the CNFET SRAM bit cells**

Since the formation of the CNTs is dependent on various factors like temperature, catalyst impurities etc. [85]. A batch of CNT is a mixture of metallic and semiconducting CNTs. As a result, there is a process and mismatch variation in the circuit. Although the mixture of semiconducting and metallic CNTs [84] can be segregated but achieving CNTs with same chirality is challenging. To verify the proposed SRAM bit cell we performed the monte-carlo process and mismatch variation for the RSNM along the mean chiral value of (19,0) as shown in Figure 42. Butterfly curve shows symmetric lobes for the proposed cell that indicates the balance of load even though the circuit is asymmetrical.

**Figure 42: 1000 points Monte-carlo simulation of Read-SNM for the proposed SRAM bit cell**
Schmitt trigger have had been neglected in designing logic circuit and memory design due to high power consumption and delay time. But it provides an immunity against process variation, that prime to a reliable circuit at low voltages. Figure 43 (a) shows the 6σ data distribution for write operation ‘1’ for proposed CNFET SRAM bit cell, it showed a mean 7.6ps which is than the conventional bit cells (Figure 43 (b)). This higher write time is associated to the single node charging and discharging during the operation.

Average power for read and write operation of the proposed bit cell is 408nW with a standard deviation of 217nW, which is better compared to the conventional 8T bit cell which has a average power mean of 2.8µW and standard deviation of 694nW across the 6σ gaussian distribution of the chirality. However, conventional 6T bit cell had 33.3% less average power but 50nW more deviation. This is due to the presence of Schmitt trigger. Figure 44 (a) shows the distribution of the average power for all the bit cells.

Figure 44 (b) shows the hold power distribution for 6T, 8T and proposed architecture, where the mean values for the hold power is 782.9n, 782.9nW and 598.9nW, respectively.
5.3.3 Layout

The layout of the proposed cell is based on DEP process. DEP is a prominent technique for placing large number of CNT over a larger area [86], and can assist in achieving a two metal layer SRAM bit cell design. When CNT solution is dispensed over two electrodes with radio frequency source attached to it, CNTs forms an electric field around themselves which helps in the alignment of CNTs [87] between the electrodes. So, for the purpose of only placing and aligning the CNTs first mask is designed. Metal 1 can be etched after the DEP, however we used the metal for routing .

Figure 45 and Figure 46 shows two masks for the layout. Metal 1 which is designed for DEP, also forms the routing for the WWL and RWL which covers a large die traces on SRAM. After the deposition of metal 1 CNTs can be placed using DEP.
The second step is the etching of unwanted area. This step helps in removing the unwanted metal which was required for DEP but not in the memory bit cell. It can also be used for removing the unwanted CNTs [88].

Figure 46, shows the final layout of the proposed SRAM showing metal 2 which is used for forming the electrode for the source and drain.
5.3.4 CNFET based in-memory computing

Since memory has a high internal bandwidth, new memory architecture is gaining popularity within the SoCs, it will be useful in designing single level setting for a monolithic memory [89]. These modifications of the conventional memory provide a simple interconnection between operational processing elements and the memory which increases the computational capability inside the chip, without any major modification. However, some of the challenges may arise for the programmers. For example, new data structure for high performance CRAM which will be better than parallel data structure on multi-core machines[66].

The architecture of the in-computing SRAM based processing is conceptualized on the process of read, compute and write [90] which has been proposed using the MOS technology with single ended read operation. But in this work, it has been further improved to assist the single-ended SRAM read as well as write operation. Moreover, the earlier proposed architecture’s logical circuit during the XOR operation uses the cell ground boost technique [91] which is although categorized as read assist technique can be a disadvantage for the half selected bit cell by reducing their hold SNM. For developing in-memory logical computation, we take the advantage of the single-ended operation of our proposed SRAM to develop a circuit capable of reading and implementing NAND, NOR and XOR logical operation on the stored bit. Additionally, the read and write port of the proposed SRAM is decoupled, which is an advantage for synchronous operation on the target bit cell [4]. Figure 47 (a) shows the overall architecture for implementing the logical operation and Figure 47 (b) shows inside of the logical circuit. Logical implementation of the circuit is designed based upon the voltage division principle and utilizing Schmitt trigger for writing and reading at multiple voltages. The circuit for the logical operation is an additional peripheral circuit, that can be placed between the SRAM bit cell and the single ended sense amplifier of one column [40]. SRAM bit cell and peripheral circuit is designed in such a way so that the overall SRAM architecture does not change.
Logical NOR

For storing the value and implementing NOR operation, the WRITE operation remains the same for both of the cells. Logical NOR operation is based on the negation of two parallel switches (Table 2). Therefore, the output for NOR operation is 1, only when both the switches are OFF.

Table 2: Truth table for NOR gate

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
Now consider two SRAM memory cell from the column vector which will be input for the NOR operation in the memory. During the NOR operation, the RBL is pulled up to VDD and the logic circuit gives bit 1 at the NAND/NOR pin, shown in Figure 47 (a). The inverters are designed with their threshold near to VDD/2 by placing same number of CNTs in pCNFET and nCNFET. The transient simulation (Figure 48) shows 3 cases: A = 1 B = 0; A = 1 B = 1; and A = 0 B = 0 respectively. Before every read operation the bits are stored with a value. For the NOR operation, the ordered read conditions are: RWL = 1 and RBL = 1. When the stored bit is zero, there is a drop in RBL, this shows that only when both the stored bits are 1. Since the drop created in 1/0 and 0/1 is less than 0.5V, an inverter is used for switching the signal to give the desired result. Therefore, by activating RWL and RBL we can observe NOR function. Figure 48 shows the transient response of the wired NOR circuit, where the NOR operation is performed only when RWL is activated.

![Figure 48: NOR operation using the proposed SRAM using the in-memory computing configuration](image)

The first two pulses represent the input ‘1/0’ and ‘1/1’ from the memory cell A/B, and the output we get is 0. Whereas the last pulse represents the case ‘0/0’ and the output we get from the pin is 1

*Logical NAND*

NAND gate output is only 0 when either the inputs are 1, or it can be expressed as the negation of two switches in series (Table 3). When RWL is activated, RBL is generally charged to VDD and then pulled down if any of the stored bit is 0.
Some of the previous approaches including ours [58], designed an inverter with its inverting point above the voltage drop created by the case ‘1/0’ or ‘0/1’ which can be achieved by changing the number of CNTs in nCNFET. However, increase in CNTs will increase the area and the variation in the circuit, inverting the output from askew inverter. Therefore, we applied variable voltage method, in which RBL is discharged, i.e., RBL= 0. During the NAND operation RBL is pulled down, which leads to charging of the read bit line through current in the memory cell. It is not sufficient enough to reach the switching voltage of the peripheral circuit’s inverter. Therefore, during ‘0/1’, ‘1/0’ and ‘0/0’ cases there is no transition in the peripheral circuit from 1 to 0. Whereas, in case of 1/1, the charging from the cell to the RBL is more than 0.5V which is more than the tripping point of the inverter. Hence, the circuit works as wired- NAND.

![Image](image.png)

**Figure 49: NAND operation using the proposed SRAM using the in-memory computing configuration**

Figure 49 shows transient simulation for 2-input memory cell showcasing NAND gate using in-memory computation logic circuit.
**Logical XOR**

Logical XOR operation plays an essential operation in AI application [92]. Therefore, inclusion of XOR operation becomes essential. The output for logical XOR is 1 only when the number of inputs in logic 1 are even (Table 4).

<table>
<thead>
<tr>
<th>A</th>
<th>B</th>
<th>Output</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

*Table 4: Truth table for XOR gate*

It can also be considered as a 2 input comparator which gives an output 1 only when there is a difference in the operands. Since NOR and NAND operation have the identical output for ‘1/1’ and ‘0/0’ case, a comparator can be designed using their logical combination. Therefore, combination of designed NAND and NOR can be exploited to implement the circuit for logical XOR.

The design of the proposed circuit is based on pass transistor circuit. Where the inverter P1/N1 and P2/N2 acts as the pass transistor and P3/N3 as load transistor for performing the XOR operation. For performing the XOR operation, RBL is pre-charged to VDD/2, so that the input for the peripheral circuit is close to the threshold point of the inverters present inside the logical circuit. This threshold/switching voltage is essential because an exclusive voltage drop in RBL is associated with every case. This voltage drop act as an electro-potential controlling the subband of the CNTs. Hence, controlling the output of inverters in the peripheral logical circuit. For performing the XOR, RBL is pre-charged to 0.5 V and the RWL is given logic 1. Now, let us consider the stored bits are 1 and 0 in A and B, respectively. As the operand B is storing 0 there is a discharge of the RBL voltage as read current into the corresponding memory cell. However, cell A balances the voltage drop by discharging into the RBL. As a result, RBL voltage is maintained at ~0.47 V. Subsequently, this voltage drop act as an input for the peripheral circuit. For unmatched operands, the RBL voltage drops down to 0.47V, which is low enough for decreasing the subband gap in P1 and P2, and high enough for increasing the subband gap in N1 and N2. Resulting in logic 1 input for the load inverter which switches ON the N3 that is connected to P1/N1 at the input and P2/N2 at the output.
The logic 0 is generated due to the dominance of N3 over P1, since the P1 transistor is too low in the linear region. However, when $A = 0$ and $B = 0$ the RBL is more discharged through the memory cell, reducing the RBL equal to 0.06V. Due to which the electro-potential across the gate shifts the subband much closer for pCNFET this brings P1 and P2 nearer to saturation regime. As a result P1 dominates over N3 and generates a logic 1 around 0.6V which can be directly used for writing in the proposed memory cell, shown in figure 50. However, with low voltages direct write operation might face latency (Figure 37). Subsequently, when $A = 1$ and $B = 1$, the working remains the same and conversely the logic 1 generated is 0.9V which can be used directly by a memory cell for a write operation, as presented in Figure 37. During this case, RBL is charged by the memory cell from VDD/2 to 0.63V, that is sufficient for inversion at the pass transistors, but since the input is not high enough at the pass gate inverters P3 dominates over the weak N1. Figure 50 shows the transient response of the XOR operation using the proposed architecture and SRAM bit cell.
Conclusion
For the new age of AI technology, the memory architecture needs to be modified so that the internal bandwidth is fully exploited at low power consumption, and Processing in-memory can be the potential solution aimed towards this new computation architecture for emerging data sensitive application; especially using the beyond CMOS technology which have faster carrier mobility. The proposed architecture with the single-ended operation gives an optimal solution for beyond CMOS technology by lowering power consumption and write time using the CNFET Schmitt trigger. Firstly, the proposed SRAM cell showed 60% higher RSNM compared to the conventional CNFET 6T SRAM bit cell and 8% lower RSNM compared to the conventional CNFET 8T SRAM whereas the proposed SRAM had 83% higher WSNM compared to both of the conventional SRAM. Secondly, the dynamic signal based Schmitt trigger showed a significant low read current with a lower variance over the 6σ range (Gaussian distributed Monte-Carlo simulation) of the CNT’s chirality compared to the conventional 8T SRAM bit cell. Lastly, the proposed SRAM showed less power consumption compared to the 8T SRAM bit cell which has same number of CNFETs. Moreover, the proposed SRAM can be read and written over multiple voltages. So it was advanced further, to show the processing in-memory concept using a peripheral circuitry which is able to perform logical NOR, NAND and XOR operations, that can be implemented for 3D monolithic memory design. The work showed NOR, NAND and XOR transient simulations for all the cases. The work also shows the layout for the proposed SRAM bit cell that can be used in future for fabrication using Dielectrophoresis.
Future Work

This work showed how CNFET can be a prospective beyond CMOS technology that can be implemented in the field of logic circuit and memory. Though according to the survey mentioned in ITRS 2017, the CNFET is not promising technology for memory but it can be correlated to the fact that there not much work has been done in this field. But with the upcoming of 3D stacking and monolithic memory CNFET’s use can see an up rise. In 3D memory, the process is limited by high temperature and mechanical process [93], therefore DEP can be a useful technique for placing the CNTs. And processing in-memory can be used for performing the parallel computation with the microprocessor, which can be silicon based.

Although the work showed the experimental procedure for the DEP, the process can be further enhanced by, firstly, accurately measuring the weight of the CNTs which will lead to better known concentration, secondly, the gate interaction from the bottom gate is not effective in manipulating the electro-potential-subband interaction which can be improved with a top gate formation, which couldn’t be achieved by us due to the lack of infrastructure. And lastly, bottom gate CNFET showed sensitivity to temperature and humidity, therefore it can be used for making sensors.

For the second part of the thesis, monte carlo simulation is done according to the Gaussian distribution which is used in the industry for conventional MOS technology, might not be suitable for the CNTs chirality simulations. The next best fit according to the literature is lognormal distribution, whose application in the statics file is currently limited by Cadence Virtuoso. Simulations are done using ideal condition, no parasitic or load has been attached which can be done to validate before fabrication. Also, simulation is done in picoseconds due to the server limitation. The proposed architecture of in-processing memory can be useful for performing parallel computing within the 3D monolithic memories in future.
References


[38] P. Lamberti and V. Tucci, “Impact of the Variability of the Process Parameters on CNT-Based
Nanointerconnects Performances: A Comparison Between SWCNTs Bundles and MWCNT,”


Appendix

Code for Band diagram of graphene

```matlab
x=linspace(-2.5,2.5,50); y=linspace(-9,10,20);
[a,b]=meshgrid(x,y)
E_bonding= 3.1*sqrt(1+4*cos(b*0.1*pi).*cos(a*1.2282)+4*(cos(a*1.2282)).^2);
E_antibonding=-
3.1*sqrt(1+4*cos(b*0.1*pi).*cos(a*1.2282)+4*(cos(a*1.2282)).^2);
figure
surf (a,b, E_bonding)
title('Energy dispersion relation of graphene')
xlabel('K(1/A)')
ylabel(' Energy of graphene(ev)')
hold on
surf(a,b, E_antibonding)
figure
plot(x, E_bonding)
title('Energy dispersion relation of graphene')
xlabel('K(1/A)')
ylabel(' Energy of graphene(ev)')
hold on
plot(x, E_antibonding)
hold on
plot(x, E_antibonding)
hold on
surf (a,b, E_bonding)
hold on
surf (a,b,zz)
```
Raman Spectrum

![Raman Spectrum Diagram](image)

Figure 51: Raman spectroscopy for the CNTs on the sample over 4 spots using 735nm near IR laser, the area between the red line shows the presence of metallic-CNTs.

Resist Recipe

For making the bi-layer resist Microchem PMMA (polymethyl methacrylate) was used.

Table 5 Resist layer 1 recipe

<table>
<thead>
<tr>
<th>Step</th>
<th>RPM</th>
<th>Time (sec)</th>
<th>Acceleration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000</td>
<td>5</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>2000</td>
<td>5</td>
<td>400</td>
</tr>
<tr>
<td>3</td>
<td>2000</td>
<td>90</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>5</td>
<td>400</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Layer 1</th>
<th>Resist Layer:</th>
<th>PMMA 495 A6</th>
<th>Bake time @temperature:</th>
<th>50 min @200°C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Saturation Time (Secs)</td>
<td>60</td>
<td>Thickness achieved</td>
<td>330-340nm</td>
<td></td>
</tr>
</tbody>
</table>
Table 6 Resist layer 2 recipe

<table>
<thead>
<tr>
<th>Step</th>
<th>RPM</th>
<th>Time (sec)</th>
<th>Acceleration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1000</td>
<td>5</td>
<td>200</td>
</tr>
<tr>
<td>2</td>
<td>1000</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>4000</td>
<td>30</td>
<td>500</td>
</tr>
<tr>
<td>4</td>
<td>4000</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>10</td>
<td>50</td>
</tr>
</tbody>
</table>

Saturation Time (Secs) | 0 |Thickness achieved | 360-375nm | Bake time@temperature: | 50 min @200°C |

Resist Layer: PMMA 950 A4

CNT other solvent choices

Figure 52: Raymore CNT in Toluene

The evaporation temperature for the Toluene was too short for effective use of DEP, since the drop would evaporate as soon as it touches the surface.
IPA had a moderate evaporation temperature but the dissolution of the CNTs was not enough, as it had more patches than single wires.

Parameter measurement for curve fitting

Figure 53: CNTs in (Iso-propanol Alcohol) IPA

Figure 54: Capacitance Measurement of the silicon wafer.
Figure 55: Refractive index of the oxide used v/s the incident wavelength using Ellipsometry

Images from other sample with different parameters

Figure 56: SEM for the electrode with DEP parameters $2V_{pp}, 150\, kHz$ and 6 secs
Figure 57: SEM for the electrode with DEP parameters $2 \text{ V}_{pp}$, 150 kHz and 10 secs

AFM

Figure 58: AFM for the electrode with DEP parameters $2 \text{ V}_{pp}$, 150 kHz and 3 secs
Figure 59: AFM for the electrode with DEP parameters $2V_{pp}, 150 \text{ kHz}$ and 10 secs

Figure 60: AFM for the electrode with DEP parameters $2V_{pp}, 150 \text{ kHz}$ and 20 secs
Monte-carlo statics variation model file

parameters monten1 = 0 monten2 = 0

statistics {
    mismatch {
        vary monten1 dist=gauss std=3
        vary monten2 dist=gauss std=3
    }
}

process {
    vary monten1 dist=gauss std=3
    vary monten2 dist=gauss std=3
}
}