Neuromagnetic Fields and Brain-Inspired Hybrid Analog-Digital Computation

by

Vivek Subramanian

Department of Biomedical Engineering
Duke University

Date: ______________________

Approved:

__________________________
Miguel A. L. Nicolelis, Supervisor

__________________________
Craig Henriquez

__________________________
Mikhail Lebedev

__________________________
Allen Song

__________________________
Trong-Kha Truong

Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering in the Graduate School of Duke University 2018
ABSTRACT

Neuromagnetic Fields and Brain-Inspired Hybrid Analog-Digital Computation

by

Vivek Subramanian

Department of Biomedical Engineering
Duke University

Date: ____________________

Approved:

___________________________
Miguel A. L. Nicolelis, Supervisor

___________________________
Craig Henriquez

___________________________
Mikhail Lebedev

___________________________
Allen Song

___________________________
Trong-Kha Truong

An abstract of a dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biomedical Engineering in the Graduate School of Duke University 2018
Abstract

Brain-inspired computing architectures such as neural networks and neuromorphic chips have demonstrated promise in performing complex pattern recognition tasks by coarsely mimicking synaptic activity in software and hardware. In this dissertation, we take a departure from these more traditional methods which are confined by what we know about the dynamics of synaptic computation and introduce a brain-inspired hybrid analog-digital computing paradigm involving magnetic fields. We first review biomagnetic fields - a wide array of topics is covered to spark the interest of the reader in the field of neuro-biomagnetism and to provide a general overview of the field that explains (1) various techniques to measure, quantify, and model the magnetic signals generated by neurons; (2) how magnetic stimulation can affect neurons; and (3) the clinical relevance of these findings. These highlight the importance of magnetism in biology and neural signal processing and provide motivation for engineering magnetically-based computational devices. We then introduce a new hybrid analog-digital computing device inspired by the interplay between neural activity and its induced magnetic fields. We show that magnetic fields can interact nonlinearly in analog in a ferromagnetic medium. Specifically, the magnetic flux induced by two alternating magnetic fields can be employed to perform an absolute difference, or smooth XOR, operation. The physical structure of the analog device is based on a white matter tractography analysis; hence, we call it the neuromagnetic reactor. We also describe our design of a scalable implementation of a perceptron in hardware,
which provides a digital 0-1 output. We demonstrate in a synthetic environment that these two systems together allow an organism to learn from and react appropriately to its environment. Although the design presented here is a proof-of-concept, it can be improved to yield not only new ways to study brain function but also new brain-inspired computing architectures based on magnetic fields.
To my loving parents, Mohan and Viji Subramanian.
Contents

Abstract iv
List of Tables x
List of Figures xi
List of Abbreviations and Symbols xix
Acknowledgements xxii

1 Introduction 1
   1.1 Preface . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
   1.2 Outline . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2

2 Neuromagnetic Fields 4
   2.1 Introduction . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 4
   2.2 Overview of key magnetic sensing and imaging modalities . . . . . . 6
      2.2.1 Superconducting quantum interference devices (SQUIDs) and
            magnetoencephalography (MEG) . . . . . . . . . . . . . . . . . . . . 6
      2.2.2 Magnetic resonance imaging (MRI) . . . . . . . . . . . . . . . . 8
      2.2.3 Functional magnetic resonance imaging (fMRI) . . . . . . . . . 9
      2.2.4 Neuronal current MRI (nc-MRI) . . . . . . . . . . . . . . . . . . 10
   2.3 Quantifying neuromagnetic signals . . . . . . . . . . . . . . . . . . . 11
      2.3.1 Individual axons . . . . . . . . . . . . . . . . . . . . . . . . . . 11
      2.3.2 Nerves . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.3 Whole brain and spinal cord</td>
<td>22</td>
</tr>
<tr>
<td>2.3.4 Promising new magnetic sensing / imaging technology</td>
<td>31</td>
</tr>
<tr>
<td>2.4 Non-synaptic interaction among neurons</td>
<td>33</td>
</tr>
<tr>
<td>2.4.1 Electric</td>
<td>33</td>
</tr>
<tr>
<td>2.4.2 Magnetic</td>
<td>36</td>
</tr>
<tr>
<td>2.5 Magnetic stimulation of neuronal activity</td>
<td>39</td>
</tr>
<tr>
<td>2.5.1 TMS</td>
<td>39</td>
</tr>
<tr>
<td>2.5.2 Interaction with ambient magnetic fields</td>
<td>43</td>
</tr>
<tr>
<td>2.6 Biomagnetite and its implications on memory and disease</td>
<td>45</td>
</tr>
<tr>
<td>2.6.1 Presence in non-human organisms</td>
<td>45</td>
</tr>
<tr>
<td>2.6.2 Pathological magnetite and neurodegenerative disorders</td>
<td>46</td>
</tr>
<tr>
<td>2.6.3 Potential for memory storage</td>
<td>48</td>
</tr>
<tr>
<td>3 Brain-inspired and hybrid analog-digital computing</td>
<td>50</td>
</tr>
<tr>
<td>3.1 Neural networks and neuromorphic computing</td>
<td>51</td>
</tr>
<tr>
<td>3.1.1 Historical background</td>
<td>51</td>
</tr>
<tr>
<td>3.1.2 Architectures</td>
<td>54</td>
</tr>
<tr>
<td>3.1.3 Implementations</td>
<td>59</td>
</tr>
<tr>
<td>3.2 Hybrid analog-digital theory of neural computation</td>
<td>64</td>
</tr>
<tr>
<td>3.2.1 The Relativistic Brain</td>
<td>65</td>
</tr>
<tr>
<td>3.2.2 Hypothesis</td>
<td>66</td>
</tr>
<tr>
<td>4 Analog computation: the neuromagnetic reactor</td>
<td>68</td>
</tr>
<tr>
<td>4.1 Diffusion weighted imaging</td>
<td>69</td>
</tr>
<tr>
<td>4.2 Tractography</td>
<td>71</td>
</tr>
<tr>
<td>4.3 Large-scale endogenous cortical magnetic and electric fields</td>
<td>74</td>
</tr>
<tr>
<td>4.4 Neuromagnetic reactor</td>
<td>78</td>
</tr>
</tbody>
</table>
4.4.1 Design 1: wire model ........................................... 78
4.4.2 Design 2: 3D printout ......................................... 96
4.4.3 Design 3: wire model with coils ............................. 105
4.4.4 Nonlinear interaction of magnetic fields .................... 132
4.4.5 Design 4: system of Nanoperm toroidal cores .......... 157

5 Digital computation: the perceptron .......................... 159
  5.1 The perceptron learning algorithm ............................. 159
  5.2 Hardware implementation of perceptron ....................... 163
    5.2.1 Timing circuit ............................................. 164
    5.2.2 Data input and prediction .................................. 170
    5.2.3 Phase identifier .......................................... 170
    5.2.4 Learning circuitry ......................................... 175
  5.3 LTspice simulations ........................................... 179
  5.4 Physical implementation: design choices ..................... 190
    5.4.1 Design choices and KiCad schematics .................... 190

6 Navigating an environment ...................................... 198
  6.1 Analog ....................................................... 199
  6.2 Digital ....................................................... 206
  6.3 Hybrid ....................................................... 207

7 Conclusions and future work .................................. 218

A Magnetic field and hysteresis measurements ................. 230

Bibliography .......................................................... 244

Biography .............................................................. 260
List of Tables

4.1 Amplitudes of induced magnetic flux densities in ferrofluid at fundamental and harmonic frequencies due to 50 Hz input at 173 A/cm... 129
4.2 Amplitudes of induced magnetic flux densities in ferrofluid at fundamental and harmonic frequencies due to 50 Hz input at 345.6 A/cm... 130
4.3 Amplitudes of induced magnetic flux densities in ferrofluid at fundamental and harmonic frequencies due to 93 Hz input at 140 A/cm... 131
4.4 Flux densities at difference frequencies with and without static magnetic field. ................................. 154
4.5 Truth table for exclusive-OR (XOR) operator. ......................... 154
6.1 Resulting difference frequency $d$ for the first ten trials of the reactor experiment. ................................. 213
## List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Diagram of a standard feedforward neural network with an input layer, two hidden layers, and an output layer.</td>
</tr>
<tr>
<td>3.2</td>
<td>CNN architecture.</td>
</tr>
<tr>
<td>3.3</td>
<td>Block diagram of a RNN with one hidden layer (Britz, 2015).</td>
</tr>
<tr>
<td>4.1</td>
<td>Diffusion tensor images.</td>
</tr>
<tr>
<td>4.2</td>
<td>Output of TRACULA.</td>
</tr>
<tr>
<td>4.3</td>
<td>Design 1: wire model of neuromagnetic reactor.</td>
</tr>
<tr>
<td>4.4</td>
<td>Experimental setup to measure magnetic field of a single wire and a pair of wires in various media.</td>
</tr>
<tr>
<td>4.5</td>
<td>Results of single wire experiment in air with blue wire driven at 50 Hz with the sensor placed on the plane between the two wires.</td>
</tr>
<tr>
<td>4.6</td>
<td>Results of the same experiment but with green wire driven at 50 Hz.</td>
</tr>
<tr>
<td>4.7</td>
<td>Results of two wire experiment performed in air with both wires driven at 50 Hz with the same magnetic field strengths.</td>
</tr>
<tr>
<td>4.8</td>
<td>Results of single wire experiment performed in air with green wire driven at 93 Hz.</td>
</tr>
<tr>
<td>4.9</td>
<td>Results of two wire experiment performed in air with the blue wire driven at 50 Hz and the green wire driven at 93 Hz.</td>
</tr>
<tr>
<td>4.10</td>
<td>Example of skin effect in a conductor.</td>
</tr>
<tr>
<td>4.11</td>
<td>Results of single wire experiment in saline with blue wire driven at 50 Hz with the sensor placed on the plane between the two wires.</td>
</tr>
<tr>
<td>4.12</td>
<td>Results of the same experiment but with green wire driven at 50 Hz.</td>
</tr>
</tbody>
</table>
4.13 Results of two wire experiment performed in saline with both wires driven at 50 Hz with the same magnetic field strengths. .......................... 90
4.14 Results of single wire experiment performed in saline with green wire driven at 93 Hz. ................................................................. 91
4.15 Results of two wire experiment performed in saline with the blue wire driven at 50 Hz and the green wire driven at 93 Hz. ...................... 92
4.16 Results of single wire experiment in ferrofluid with blue wire driven at 50 Hz with the sensor placed on the plane between the two wires. . 93
4.17 Results of the same experiment but with green wire driven at 50 Hz. 94
4.18 Results of two wire experiment performed in ferrofluid with both wires driven at 50 Hz. ................................................................. 95
4.19 Results of single wire experiment performed in ferrofluid with green wire driven at 93 Hz. ................................................................. 96
4.20 Results of two wire experiment performed in ferrofluid with the blue wire driven at 50 Hz and the green wire driven at 93 Hz. ............... 97
4.21 Monte Carlo distribution of pathways for the forceps major for the single, healthy subject. ................................................................. 98
4.22 Means of each of the 18 tracts and along with one standard deviation in the x-, y-, and z-directions. ...................................................... 99
4.23 Processing steps to produce tubular structures. ............................. 101
4.24 Orthographic views of simple model with circular cross-sections. ... 102
4.25 Isometric view of the simple model (with circular cross-sections). ... 103
4.26 3D printout of simple model. Colors correspond to those in figure 4.2c. 104
4.27 Views of the surfaces and solid of the forceps major created from point clouds in Solidworks for the model with elliptical cross-sections. ... 106
4.28 Orthographic views of Solidworks designs for the realistic model with elliptical cross-sections. ......................................................... 107
4.29 Isometric view of the realistic model with elliptical cross-sections. ... 108
4.30 Additional views of the surfaces of the realistic model. .................... 109
4.31 Design 3: wire model of neuromagnetic reactor with twelve coils. ... 110
4.32 The numeric and color code shown here indicates the tracts from figure 4.2 to which each of the coils in this reactor design correspond.

4.33 Experimental setup and circuit diagram from coil experiment.

4.34 Result of coil experiment in air in which a 50 Hz, 16.9 A current was passed for four seconds through a 100 turn, 3 in. diameter coil.

4.35 Result of coil experiment in air in which a 50 Hz current of increased amplitude 33.3 A was passed for four seconds through the same coil.

4.36 Result of coil experiment in air in which a 93 Hz current of amplitude 14.2 A was passed for four seconds through the same coil.

4.37 Result of the coil experiment in air in which both 50 and 93 Hz sinusoids were inputted into the coil for four seconds.

4.38 Result of coil experiment in saline in which a 50 Hz, 16.5 A current was passed for four seconds through a 100 turn, 3 in. diameter coil.

4.39 Result of coil experiment in saline in which a 50 Hz current of amplitude 32.71 A was passed for four seconds through the same coil.

4.40 Result of coil experiment in saline in which a 93 Hz current of amplitude 14.0 A was passed for four seconds through the same coil.

4.41 Result of the coil experiment in saline in which both 50 and 93 Hz sinusoids were inputted into the coil for four seconds.

4.42 Amplitude spectrum of magnetic field strength when 50 Hz, 16.3 A sinusoid was applied to a coil inside which sat a ferrofluid medium.

4.43 Result of coil experiment in ferrofluid in which a 50 Hz, 16.3 A current was passed for four seconds through a 100 turn, 3 in. diameter coil.

4.44 Hypothetical magnetic flux density induced by a sinusoidal field strength at 5 Hz.

4.45 Amplitude spectrum of magnetic field strength when 50 Hz, 32.6 A sinusoid was applied to a coil inside which sat a ferrofluid medium.

4.46 Result of coil experiment in ferrofluid in which a 50 Hz current of amplitude 32.6 A was passed for four seconds through the same coil.

4.47 Result of coil experiment in ferrofluid in which a 93 Hz current of amplitude 13.2 A was passed for four seconds through the same coil.
4.48 Result of the coil experiment in ferrofluid in which both 50 and 93 Hz sinusoids were inputted into the coil for four seconds. 127

4.49 Nanoperm toroidal core. 136

4.50 Magnetic properties of Nanoperm shown in red. 137

4.51 Circuit used to measure hysteresis in Nanoperm core. 138

4.52 Circuit used to verify nonlinear interaction of alternating magnetic fields in the presence of a direct magnetic field. 140

4.53 Nanoperm hysteresis loops measured during single frequency inputs. 142

4.54 FFT of measured magnetic field for single frequency inputs. 143

4.55 Nanoperm hysteresis loops measured when driven by two frequencies of AC without static magnetization. 144

4.56 FFT of magnetic field when core is driven by two frequencies without static magnetization. 145

4.57 50 Hz hysteresis loops measured in the presence of direct magnetic field at various field strengths. 146

4.58 FFT of magnetic field when driven by 50 Hz at various direct magnetic field strengths. 147

4.59 100 ms snapshot of the induced flux density due to 50 Hz sinusoidal input with and without the addition of a static magnetic field. 148

4.60 Simulated flux density during saturation. 149

4.61 Nanoperm hysteresis loops measured during single frequency inputs in the presence of 2991.0 mA/cm direct magnetic field strength. 150

4.62 FFT of measured magnetic field for single frequency inputs in the presence of 2991.0 mA/cm direct magnetic field strength. 151

4.63 Nanoperm hysteresis loops measured when driven by two frequencies in the presence of static magnetic field. 152

4.64 FFT of magnetic field when core is driven by two frequencies in the presence of static magnetic field. 153

4.65 Theoretical absolute difference operations for two ranges. 155

4.66 Interacting flux densities with and without static magnetic field. 155
5.17 reset2 phase of learning circuitry. \( V(C_{\text{weight}}) \) is reset, priming it for the update. .......................................................... 188

5.18 update phase of learning circuitry. \( V(C_{\text{temp}}) \) equilibrates with \( V(C_{\text{weight}}) \), completing the PLA update. ......................................................... 189

5.19 Long-term behavior of \( V_w \), showing that the weight converges to a reasonable value based on the data that were inputted. .......................... 190

5.20 Assembled printed circuit board (PCB) showing the design choices adopted for the physical design of the perceptron. ............................... 191

5.21 PCB schematic for perceptron: top level overview. .............................. 192

5.22 PCB schematic for perceptron: timing circuit. ..................................... 193

5.23 PCB schematic for perceptron: data input and prediction circuit and phase identifier. ................................................................. 195

5.24 PCB schematic for perceptron: learning circuitry. ............................... 196

6.1 Block diagram showing the three layers of the neuromagnetic reactor. 204

6.2 \( \bar{b}^a \) and corresponding \( \bar{c}^a \) values for all possible values of elements \( a \). 205

6.3 Spectrograms for the three Nanoperm cores in the first layer of the reactor. ................................................................. 208

6.4 Spectrograms for the three Nanoperm cores in the second layer of the reactor. ................................................................. 209

6.5 Spectrograms for the final Nanoperm cores in the third layer of the reactor. ................................................................. 210

6.6 Plot of the time series of \( x \). ............................................................. 211

6.7 Plot of the values of \( d \) corresponding to \( x \) from figure 6.7. .................. 212

6.8 Plot of the time series of \( y \). ............................................................. 214

6.9 Plot of the output of the phase identifier window comparators for a single charge-discharge cycle of \( C_{\text{timing}} \). ................................. 215

6.10 Output of the logic gates to which the phase identifier window comparators are connected. ................................................................. 216

6.11 Plot showing the voltage \( V_w \) on the weight-storing capacitor. ............ 217

xvi
7.1 Local tractography analysis and 3D printout.

7.2 Simulation of spikes and induced magnetic fields interacting and encoding information by oscillations.

7.3 Memory circuit in which the measurements made on each of the toroids are filtered and fed back to each other.

7.4 Circuit in which three cores interact with nonlinear feedback, possibly resulting in chaotic behavior.

A.1 51 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.

A.2 67 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.

A.3 95 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.

A.4 102 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.

A.5 127 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.

A.6 147 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.

A.7 150 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.

A.8 Nanoperm hysteresis loops measured when driven by two frequencies in the presence of 59.8 mA/cm static magnetic field.

A.9 FFT of magnetic field when core is driven by two frequencies in the presence of 59.8 mA/cm static magnetic field.

A.10 Nanoperm hysteresis loops measured when driven by two frequencies in the presence of 299.1 mA/cm static magnetic field.

A.11 FFT of magnetic field when core is driven by two frequencies in the presence of 299.1 mA/cm static magnetic field.

A.12 Nanoperm hysteresis loops measured when driven by two frequencies in the presence of 598.2 mA/cm static magnetic field.
A.13 FFT of magnetic field when core is driven by two frequencies in the presence of 598.2 mA/cm static magnetic field. . . . . . . . . . . . . . . 243
# List of Abbreviations and Symbols

## Symbols

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x$</td>
<td>scalar</td>
</tr>
<tr>
<td>$\mathbf{X}$</td>
<td>vector or matrix</td>
</tr>
<tr>
<td>$\mathbb{X}$</td>
<td>vector space</td>
</tr>
</tbody>
</table>

## Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Ampere (unit)</td>
</tr>
<tr>
<td>AC</td>
<td>alternating current</td>
</tr>
<tr>
<td>AD</td>
<td>Alzheimer’s disease</td>
</tr>
<tr>
<td>AI</td>
<td>artificial intelligence</td>
</tr>
<tr>
<td>ANN</td>
<td>artificial neural network</td>
</tr>
<tr>
<td>BEDPOST</td>
<td>Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques</td>
</tr>
<tr>
<td>BOLD</td>
<td>blood-oxygen level dependent</td>
</tr>
<tr>
<td>CMOS</td>
<td>complementary metal-oxide semiconductor</td>
</tr>
<tr>
<td>CNN</td>
<td>convolutional neural network</td>
</tr>
<tr>
<td>DAQ</td>
<td>data acquisition</td>
</tr>
<tr>
<td>DC</td>
<td>direct current</td>
</tr>
<tr>
<td>DTI</td>
<td>diffusion tensor image or imaging</td>
</tr>
<tr>
<td>DWI</td>
<td>diffusion weighted image or imaging</td>
</tr>
<tr>
<td>EEG</td>
<td>electroencephalo-graphy or -gram</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>-------------</td>
<td>------------</td>
</tr>
<tr>
<td>EMG</td>
<td>electromyography or -gram</td>
</tr>
<tr>
<td>EPI</td>
<td>echo-planar imaging</td>
</tr>
<tr>
<td>FFT</td>
<td>fast Fourier transform</td>
</tr>
<tr>
<td>fMRI</td>
<td>functional MRI</td>
</tr>
<tr>
<td>FPGA</td>
<td>field programmable gate array</td>
</tr>
<tr>
<td>GPU</td>
<td>graphics processing unit</td>
</tr>
<tr>
<td>LGF</td>
<td>local geomagnetic field</td>
</tr>
<tr>
<td>LSTM</td>
<td>long short-term memory (RNN)</td>
</tr>
<tr>
<td>LTU</td>
<td>linear threshold unit</td>
</tr>
<tr>
<td>MCMC</td>
<td>Markov chain Monte Carlo</td>
</tr>
<tr>
<td>MDL-PFC</td>
<td>mid-dorsolateral prefrontal cortex</td>
</tr>
<tr>
<td>MEG</td>
<td>magnetoencephalo-graphy or -gram</td>
</tr>
<tr>
<td>MNG</td>
<td>magnetoneurography</td>
</tr>
<tr>
<td>MRI</td>
<td>magnetic resonance image or imaging</td>
</tr>
<tr>
<td>nc-MRI</td>
<td>neuronal current MRI</td>
</tr>
<tr>
<td>NEMF</td>
<td>neuroelectromagnetic field</td>
</tr>
<tr>
<td>NI</td>
<td>National Instruments</td>
</tr>
<tr>
<td>NV</td>
<td>nitrogen-vacancy</td>
</tr>
<tr>
<td>Ω</td>
<td>ohms (unit)</td>
</tr>
<tr>
<td>op amp</td>
<td>operational amplifier</td>
</tr>
<tr>
<td>PCB</td>
<td>printed circuit board</td>
</tr>
<tr>
<td>PET</td>
<td>positron emission tomography</td>
</tr>
<tr>
<td>PLA</td>
<td>perceptron learning algorithm</td>
</tr>
<tr>
<td>PSP</td>
<td>postsynaptic potential</td>
</tr>
<tr>
<td>RBT</td>
<td>relativistic brain theory</td>
</tr>
<tr>
<td>ReLU</td>
<td>rectified linear unit</td>
</tr>
</tbody>
</table>
RF  radio frequency
RNN  recurrent neural network
rTMS  repetitive TMS
SE  spin-echo
SERF  spin-exchange relaxation-free
SQUID  superconducting quantum interference device
ssVEP  steady state VEP
sTMS  short TMS
T  Tesla (unit)
TMS  transcranial magnetic stimulation
TRACULA  TRActs Constrained by UnderLying Anatomy
V_{pp}  volts peak-to-peak
VEP  visual evoked potential
VLSI  very large-scale integration
W  watt (unit)
Acknowledgements

There are a number of people to whom I owe a great thanks for helping me through the process of obtaining this PhD. First and foremost, I would like to thank my advisor, Miguel Nicolelis, whose leadership and support were critical to my success as a graduate student. Often, the path was unclear, but I was very fortunate to have such a visionary leader guiding me. Additionally, I would like thank my remaining committee members Craig Henriquez, Mikhail (Misha) Lebedev, Allen Song, and Trong-Kha Truong for always being approachable and providing invaluable advice throughout the research process. I would especially like to thank Craig with whom I had many meaningful discussions which shaped the course of my PhD and the contents of my dissertation.

Next, I would like to thank our lab members. First, I thank the administrators including Susan Halkiotis, Terry Jones, and Laura Oliveira for their tremendous support. Their help was essential to keep the lab running smoothly, allowing me to fully focus on my research. I would also like to thank my fellow labmates (past and present) including Je Hi An, Billy Byun, Tim Hanson, Will Huffman, Peter Ifft, Joshua Khani, Misha Lebedev, Gary Lehew, Shuangyan Li, Zheng Li, Jim Meloy, Adarsh Nellore, Joey O’Doherty, Jaehong Park, Tamara Phillips, Helen Powell, Shankari Rajangam, Arjun Ramakrishnan, David Schwartz, Angelo Takigami, Paul Thompson, Eric Thomson, Po-He Tseng, Tristan Vouga, Tobias Woldt, Youlu (Lee) Xing, Amol Yadav, Allen Yin, and Katie Zhuang. I am fortunate to have had
many great discussions with all of you that have made my PhD a fruitful experience full of fun and adventure.

I would especially like to mention the following lab members. Zheng: thank you for your mentorship and guidance before and during my PhD and for priming my interest in computational research with your contributions. Joey and Tim: thank you for being available to talk with me whenever I needed your guidance. Je Hi, David, Tristan, Allen, Billy, and Will: thank you for being great officemates and for keeping me company during my days in the monkey lab. It was always a pleasure being around all of you. Misha: thank you for always being positive and upbeat and for helping come up with ideas for new analyses and methods that I could pursue in my research. Billy, Will, Jaehong, Arjun, David, Eric, Po-He, Tristan, Tobias, Lee, Amol, and Allen: thank you all so much for getting me un-stuck time and again by getting me to think outside the box. Helen, Angelo, and Adarsh: it was a pleasure mentoring you and I enjoyed watching you develop your interest and expertise in programming and data analysis.

I would also especially like to thank Gary Lehew and Jim Meloy. Thanks to both of you for being such great teachers and for always being willing to lend a hand. Jim, you have always inspired me with your patience, organization, and efficiency, which are all second to none. Gary, more than just your amazing technical skills, which many times helped me out of difficult situations, I am incredibly thankful for your words of inspiration when I felt lost or confused. Thanks for always being so positive and helping me stay positive and relaxed even when things could get stressful. There is no way I could have done this without you.

I would also like to say a special thanks to my friends Billy Byun, Peter Dong, Justin Ko, Jonathan Lee, and Kesh Rao. Billy, it’s been really great learning from and laughing with you over the past several years. I will never forget our late nights in the monkey lab. Peter, thanks for doing such a great job keeping in touch and
reminding me not to forget old friends. Justin, I know I can always count on you if I need a good laugh. Jonathan, thanks for being a great roommate and for the many great discussions, both research-related and otherwise, we’ve had over the years. Kesh, thank you for your wonderful sense of humor and the endless puns. There’s never a dull moment when you’re around.

I also would like to thank Ahmed Jaraki, whose friendship I am eternally grateful for. Ahmed, the two of us have been there for each other through difficult times, and I am so fortunate to have had your support. Thank you for listening to me when I needed someone to lend an ear and for being someone I know I could always depend on. No matter where you end up in the coming months, I am confident ours is a friendship that no amount of distance can separate.

To Charmaine Mutucumarana, I am very grateful for your support of my endeavors and all the great memories of time spent together outside of the lab. It has been fun goofing off with you while watching movies and playing video games; cooking some exotic recipes like lumpy lumpias; and traveling to the tea plantations of Sri Lanka (wait a second...!). I have always admired your incredible work ethic and dedication to your profession: you have an immense drive to succeed and to bring those around you up with you. Thank you for your friendship.

Richard Bouchard, thank you for inspiring me to pursue a PhD. If it were not for your taking me under your wing as a “Young Padawan” when I was an undergraduate, I would never have become a “Jedi.” I am so grateful for all the times you spent talking with me, emphasizing the importance of balance between work and play. You showed me how to unlock my full potential and helped me develop the boundless work ethic that I carry with me to this day.

My uncles Suresh and Mahesh Venkatraman, thank you so much for always being available to listen to me and to discuss ideas with me whenever I needed your help. Your guidance was instrumental to my success, and I’m very fortunate to have had

xxiv
your support and encouragement whenever I was struggling. It is always a pleasure visiting you, and every time, I come back more curious and energized.

I would also like to thank my dear Menon Uncle, Jayaraman Uncle, and Padma Aunty who have been supportive of me since I was a baby. Thank you Menon Uncle for guiding young Bibek to work hard in school and to take up piano lessons at a young age. Thank you Jayaraman Uncle and Padma Aunty for supporting me since I was a baby and treating me with so much love and affection.

To my parents, you have been such an inspiration for me not only throughout my PhD but for my entire life. To my father, Mohan - who immigrated to the United States in 1974 with nothing but a bag of clothes and a mountain of loans - your courage, determination, optimism, and willingness to embrace uncertainty are a huge source of confidence for me. To my mother, Viji - who never fails to provide me with the motivation to keep working diligently - I am so thankful for your ability to make me laugh and smile. Thank you both for your support over the years. I could not have done this without you.

I also thank YouTubers CarlSagan42, Fruddle, Kyle Landry, and Smooth McGroove. While I do not know them personally, I closely followed their channels as they allowed me to relive the magic of my childhood through video game music or made me laugh with hilarious playthroughs of Mario Maker troll levels. Thank you for your effort to make such great online content that allowed me to escape the stresses of PhD life.

There are many others - fellow students, faculty at Duke and elsewhere, old friends, and family - who have defined my experience as a student from the moment I started school to the day I defended my PhD. Even if I did not mention your names here, please know that I have not forgotten you, that you were important to my success, and that you have forever shaped who I am as a student and as an individual.
1

Introduction

1.1 Preface

The human brain not only generates electrical and magnetic activity which we can record and study, it is one of the most powerful computing devices known to mankind with the capability to perform petaflops of computation with infinitesimal power consumption (Merkle, 1989). As a trained engineer and scientist-in-training, I was fascinated just as much (if not more) by the latter as the former and wondered how I could join these two seemingly disparate facts under a single hypothesis. Given that the magnitude of magnetic flux densities induced by cortical and subcortical electrical activity is relatively weak, the involvement of magnetic fields in neural computation seemed to be a bit of a stretch. However, from an engineering perspective, I realized that restricting oneself to magnetic flux densities that were of the magnitude found in the brain might close one off from potentially harnessing a powerful computing medium in magnetic fields. Hence, I developed a curiosity in the ability to compute with magnetic fields in general - even though the flux densities I was considering were outside the scope of those found in the brain. As a graduate student in the Nicolelis laboratory, I carried out independent research to test the hypothesis that we can
improve the performance (speed, efficiency, and/or accuracy) of today’s computers by exploiting hybrid architectures involving magnetic fields. My early findings and fascination with the brain inspired me to create a new computing device involving magnetic fields that loosely mimicked the machinery of the brain in hardware. With my device, I do not support or refute the claim that the brain utilizes magnetic fields to compute nor am I trying to build a brain in hardware. After all, in its current state, the device I am going to describe differs so starkly in so many ways from the human brain that I cannot come close to calling it a simulation. (To this end, state-of-the-art neuromorphic chips, which aim to model the behavior of individual synapses, are closer to a true brain in silico.) I hope my experiments and findings illustrate how hybrid systems might be useful and, at the very least, provide groundwork and inspiration for new directions in neuromorphic computing and studies of the brain’s electromagnetic fields. In this dissertation, I outline my hypotheses and findings that have stemmed from my research on neuromagnetic fields and hybrid analog-digital computing.

1.2 Outline

This dissertation is divided into six parts. In chapter 2, I first review biomagnetic fields - a wide array of topics is covered to spark the interest of the reader in the field of neuro-biomagnetism and to provide a general overview of the field that explains (1) various techniques to measure, quantify, and model the magnetic signals generated by neurons; (2) how magnetic stimulation can affect neurons; and (3) the clinical relevance of these findings. These highlight the importance of magnetism in biology and neural signal processing and provide motivation for engineering magnetically-based computational devices. This comprises the first major contribution of this thesis. Then, in chapter 3, I review the history of brain-inspired computing, with an emphasis on neural networks and hybrid analog-digital computers. This leads
into three chapters which comprise the second major contribution of the thesis - a new hybrid analog-digital computing device inspired by the interplay between neural activity and its induced magnetic fields. Chapter 4 describes experiments to verify that alternating magnetic fields interacting in analog in a ferromagnetic conductor can interact nonlinearly, producing harmonics that are not present when the fields are induced independently or together in a linear medium. The physical structure of the device is based on a white matter tractography analysis; hence, it is called the neuromagnetic reactor. Chapter 5 details my design of a scalable hardware implementation of a perceptron - a circuit which provides a digital 0-1 output by learning a threshold that separates data belonging two classes. Finally, in chapter 6, I demonstrate that these two systems can work together in a synthetic environment to allow an organism to learn from and react appropriately to its environment. I end with some concluding remarks in chapter 7, showing that although the design presented here is a proof-of-concept, it can be improved to yield not only new ways to study brain function but also new brain-inspired computing architectures based on magnetic fields.
Neuromagnetic Fields

2.1 Introduction

Biomagnetism is the study of magnetic fields originating from excitable tissue such as the heart and the brain (Hämäläinen et al., 1993). Neuro-biomagnetism, the subfield of biomagnetism dealing with neuronal tissue such as neurons and glial cells, has a long, rich history beginning with the discovery of magnetoencephalography (MEG) by David Cohen (Cohen, 1972). MEG opened doors to directly studying the magnetic field induced by neurons because of its ultra-high sensitivity. At the same time, researchers such as John Clark and Robert Plonsey were pioneering computational methods for calculating induced magnetic fields (Clark and Plonsey, 1966). These initial discoveries paved the way for a number of studies involving the measurement and modeling of magnetic fields of individual axons (Roth and Wikswo, 1985), peripheral nerve fibers (Gielen et al., 1991), and, eventually, the cortex (Yang et al., 1993) and tissues within the brain (Chow et al., 2006a).

Scientists and clinicians have aimed to understand how neuro-magnetic fields originate and how measuring these signals can be indicative of various disorders of
the nervous system. Several tools and techniques have been developed to answer questions such as, “How does the magnetic field of an axon vary as a function of space and time, and how do we model the material in and around an axon to get a sense of how the current distribution changes?”; “How can we precisely localize the cortical currents induced by sensory and/or motor stimuli via magnetic fields?”; “Can magnetic fields generated by external sources such as transcranial magnetic stimulation (TMS) coils be used for therapeutic purposes?” These range from MEG, which provides excellent spatial and temporal resolution, to magnetic resonance imaging (MRI), which is less sensitive - some would say insensitive - to the induced field but, in theory, allows characterization of the magnetic fields induced by neurons throughout the entire brain volume. New tools such as atomic magnetometers are being developed which have even higher spatial and temporal resolution than that of MEG. Here, we highlight many of the technical advances that took place in the field of neuro-biomagnetism. We recall some of the earliest experiments on individual axons and nerve fibers, and work our way through the discoveries. We consider technological innovations such as MEG and neuronal current magnetic resonance imaging (nc-MRI), a type of MRI in which magnetic fields induced by currents flowing through neurons are detected by MRI. We review theoretical breakthroughs in modeling magnetic fields of neurons, nerve fibers, and cortical tissue. We also discuss experimental results that provide a practical basis on which to ground the theory.

In addition, we also touch on some areas which are not traditionally associated with biomagnetism but which do have scientific as well as clinical relevance for researchers seeking to explain the mechanics of the nervous system and to diagnose and treat clinical conditions. The former includes an effect known as ephaptic coupling which explains the electrical coupling of neurons through non-synaptic means; the effect also has a magnetic analog which has inspired theories on how the brain stores memories and computes. Ephaptic coupling is still a fairly young discovery, and there
is still much to be learned about its potential to act at a macroscopic, whole-brain level. On the other hand, theories of magnetic coupling remain controversial and have yet to be proven or disproven by experiments. It also includes the discovery of biomagnetite in the human nervous system; while scientists have long known of the existence of biomagnetite in organisms ranging from bacteria to sharks, its discovery in humans has led to the finding that Alzheimer’s disease may be closely tied with accumulation of magnetite in the brain. We also describe TMS, a tool that allows scientists to probe the nervous system as well as administer therapy for patients.

2.2 Overview of key magnetic sensing and imaging modalities

We begin with a summary of the key sensing and imaging modalities that will be discussed.

2.2.1 Superconducting quantum interference devices (SQUIDs) and magnetoencephalography (MEG)

The superconducting quantum interference device, or SQUID, is a highly sensitive magnetic field measurement device used for measuring biomagnetic signals such as postsynaptic potentials (PSPs) and action potentials (Fagaly, 2005; Hämäläinen et al., 1993). These are two of the major types of electrical signaling in the brain. PSPs result in small deflections of approximately 10 mV in magnitude; assuming a conductivity of 1 $\Omega^{-1}m^{-1}$, a single PSP would result in a current dipole moment of about 20 fA $\cdot$ m. Action potentials, on the other hand, generate much stronger magnetic fields because the voltages and currents are much greater (100 fA $\cdot$ m as opposed to 20). A SQUID is sensitive to dipole moments of 10 nA $\cdot$ m or greater. Thus, considering PSPs only, between one and two hundred PSPs per thousand in an area of one mm$^2$ would result in a magnetic field measurable by a SQUID.

SQUIDs consist of two semicircular components of superconducting material
which are joined by two Josephson junctions (Fagaly, 2005). They come in two varieties: the radio frequency (RF) SQUID utilizes radio waves to bias the Josephson junction, and the direct current (DC) SQUID uses DC to do the same. In general, DC SQUIDs are more sensitive than RF SQUIDs and are much more commonly used. To summarize how a SQUID works, we first note that when the superconductor reaches a state in which the bias current is below a threshold, the material becomes so highly conductive that no voltage drop appears across the material. Magnetic fields generated by currents in a superconductor can be sustained without any external stimulation. By the Josephson effect, when the length of the Josephson junction is smaller than the coherence length of the superconductor, current can pass through the junction without any voltage drop. When a magnetic field is coupled into the loop by a measurement coil, the voltage across the Josephson junction changes. The voltage across the junction and the strength of the magnetic field vary periodically with a period equal to the magnetic flux quantum. By locking in the voltage at a unique point on the voltage-flux curve, one can determine the magnetic field induced.

Currents flowing between neurons produce weak magnetic fields which can be measured by a SQUID magnetometer using a technique called magnetoencephalography (MEG) (Hämäläinen et al., 1993). A SQUID is used to carry out MEG measurements because of its ability to sense very small magnetic fields such as those originating from the brain (on the order of 50 to 500 fT, about one billionth the size of the earth’s magnetic field). On the order of thousands of neurons must fire concertededly to generate a signal large enough to be measured by a SQUID.

MEG measures cortical magnetic fields that permeate radially outward from the skull (Hämäläinen et al., 1993). Endogenous currents that generate these fields flow in directions tangential to the skull, such as along fissures in the brain. Many sensory projections - including somatosensory, visual, and auditory - project to fissures in the brain, so these signals are the easiest to detect via MEG. The source currents
that generate an MEG signal (or electroencephalography, EEG, for that matter) are not unique, so a specific source model must be assumed to back-calculate the source currents. This is typically done by fitting predictions made by a model against measurements; the model with the most accurate predictions is then assumed to be a good source model for the data.

EEG is the electrical counterpart to MEG in which electric fields that originate throughout the depth of the brain are measured with electrodes at the scalp. MEG and EEG systems consist of many tradeoffs in terms of cost, ease of use, and information provided. As described by Cohen in his seminal MEG paper, MEG and EEG signals provide complementary information; MEG provides excellent spatial and temporal resolution (around a few mm and ms, respectively) while EEG provides excellent temporal resolution (around a few milliseconds) but poorer spatial resolution (Cohen, 1972). Typically, MEG measurements can be taken more quickly than EEG measurements, but they require the subject to remain stationary; on other hand, EEG allows the subject to move but the measurement apparatus is more difficult and time-consuming to set up. Cost-wise, clinical EEG systems can range in the tens of thousands of dollars while MEG scanners can range in the hundreds of thousands if not millions (Farnsworth, 2017; Wired Staff, 2007).

2.2.2 Magnetic resonance imaging (MRI)

Magnetic resonance imaging (MRI) is a magnetic contrast-based technique used to determine the composition of tissue in the body (Huettel et al., 2004). The subject is placed inside a cylindrical machine which creates a magnetic field known as \( B_0 \) between 1.5 T and 7 T with direction along the axis of the machine. This is around 10,000 times the strength of the field produced by the earth. When this field is applied, protons in the subject’s body, especially prominent in water, align with the direction of the magnetic field, either parallel or antiparallel to it. A second
field known as the magnetic gradient is used to select the slice of the subject in which imaging is performed. At the same time, a radio frequency pulse is applied to cause the protons to flip their orientation so that they are perpendicular to $B_0$. When the pulse is switched off, the protons relax back to align with $B_0$. During this process, they precess, or rotate about their axis, at a frequency known as the Larmor frequency. The stronger $B_0$ is, the greater the Larmor frequency; and the greater the Larmor frequency, the higher the resolution with which one can image the subject. When the protons precess, they give off an electromagnetic field in the form of radio-frequency photons which can be detected by pickup coils in the scanner. Different tissues have different concentrations of protons, and this provides a source of contrast in the resulting image. The signals that are measured are converted into an image via image processing algorithms, typically involving Fourier transforms. While the $B_0$ employed to perform MRI are very strong, they are generally considered to be safe because they are static fields, meaning they cannot induce voltage in the tissue.

2.2.3 Functional magnetic resonance imaging (fMRI)

Functional magnetic resonance imaging (fMRI) is a technique used to measure cortical activity by detecting the response to magnetic fields of vasculature supplying the neurons (Huettel et al., 2004). MRI and fMRI work in a similar way except for differences in how scans are acquired and the final image is reconstructed. While MRI can be performed anywhere in the body, fMRI is most often performed in the brain. In general, fMRI requires a series of MRI acquisitions to be performed. First, a high-resolution baseline acquisition is performed to determine the locations of prominent structures in the brain. Next, a series of low-resolution acquisitions are performed while the subject is performing a mental task. This provides information about how the distribution of water molecules changes while the subject is performing the task. Specifically, one is interested in the change in concentration of oxygenated blood as
the task is performed. Blood is high in water content and thus provides a prominent signal for fMRI images. Oxygenated and deoxygenated blood interact differently with the magnetic field generated by the scanner. By comparing the series of scans to the baseline, one obtains a so-called blood-oxygen level dependent, or BOLD, signal that provides information about where blood flows during task execution. The spatial resolution of fMRI is on around 3-6 mm while the temporal resolution is only on the order of seconds.

2.2.4 Neuronal current MRI (nc-MRI)

Neuronal current MRI, or nc-MRI, is a magnetic imaging technique in which neural activity is directly imaged by magnetic resonance (Hagberg et al., 2006). This is unlike EEG and MEG, which require source current models to determine where the magnetic signal was generated. Moreover, unlike MEG, nc-MRI can measure neuronal activity originating subcortically. While, in theory, one could jointly perform EEG and fMRI to achieve high spatial and temporal resolution, post-processing the measurements to integrate the information they each provide can be challenging. nc-MRI, on the other hand, has the potential to localize neural activity in three dimensions with high temporal and spatial resolution by exploiting magnetic fields generated by neurons for detection by MRI. Interactions of the fields generated by current carrying neurons with the $B_0$ of the MRI machine result in changes in the precession of the protons generating the fields. Dephasing of the precession results in a signal that can be measured by an MRI scanner pickup coil. However, as is the case for MEG, signals are very weak - in the fT and pT range - and therefore, many challenges to the utility of nc-MRI are present.
2.3 Quantifying neuromagnetic signals

In this section, we review the sources of magnetic signals generated by the nervous system and the methods employed to study them. These include individual axons, nerve bundles, and the brain and / or spinal cord as a whole. Each of these components in the hierarchy has been investigated with computational and experimental techniques. Moreover, the clinical relevance that allows these scientific discoveries to be translated into practical applications is addressed. We also briefly touch on new technology that is advancing the forefront of imaging and sensing magnetic fields produced by neural activity.

2.3.1 Individual axons

We begin by considering individual axons, which we take to be the fundamental unit of the nervous system.

*Computational*

The first models of the magnetic fields induced by changing intra- and extracellular potentials of individual axons were developed in the 1980s. Prior to the work of Swinney and Wikswo, nerves were simply treated as current-carrying wires. In their paper, the authors accounted for both the current in and around the axons and showed that magnetic field contributions are predominantly due to currents within the neuron rather than due to external or transmembrane currents (Swinney and Wikswo Jr, 1980), a finding later confirmed by Plonsey (Plonsey, 1982). Their methods involved (Clark and Plonsey, 1968; Clark et al., 1978). The authors found that the source of the magnetic field are internal axial currents and the discontinuity of the current density in the cellular boundary (Swinney and Wikswo Jr, 1980). (The conductivity of the membrane is $10^{-5}$ $\Omega^{-1}m^{-1}$ compared to that of the axoplasm which is $1 \Omega^{-1}m^{-1}$, a difference of five orders of magnitude.)
The authors applied their equations to the practical problem of solving for the magnetic field outside a crayfish lateral axon (Swinney and Wikswo Jr, 1980). The magnitude of the axial current density along the inner membrane surface, reaching a peak of about 60 A/m², was about two orders of magnitude greater than that along the external membrane surface or through the membrane, which only reached a peak of about 0.5 A/m². The resultant peak magnetic field at a distance of 120 µm from the axon due to the internal current was about 1 nT. Within 1 mm from the axon, the field strength falls as $1/\rho$ while beyond 10 mm, it falls as $1/\rho^3$, consistent with Ampere’s law and the properties of the extracellular space. The work indicated that the resulting waveform shapes were consistent with expectations based on previous measurements, but a combined experiment both measuring the magnetic field and the electric potential was not carried out.

Woosley et al. later developed a new method to calculate electrophysiological properties of an isolated, unmyelinated neuron in a uniform, conductive medium and applied it to predicting magnetic fields from action potentials traveling down a crayfish medial giant axon (Woosley et al., 1985). Only the transmembrane potential was assumed given, and the external resistance of the bath is assumed to be much less than the resistance inside the axon (the same assumption made in (Swinney and Wikswo Jr, 1980)). Near the nerve, the magnetic field is known to behave just as it does for a current-carrying wire. Far from the nerve, however, the relationship becomes more complicated because of the spatiotemporal distribution of current along the neuron. They found that the largest spatial frequency that contributes to the transmembrane potential is about $1/\Delta z$ where $\Delta z$ is the spatial width of the action potential. This implies that the membrane potential changes over large distances compared to radius of the axon. The model is insensitive to changes in external conductivity, but, in the DC limit, the internal conductivity is very influential in predicting the magnetic field. The conduction velocity is relevant to all calculations.
and sets the length scale of the action potential; moreover, the faster the action potential conduction, the smaller the current density and magnetic field. In the DC limit, the authors found that the magnetic field is proportional to the square of the axon radius. This is because the current accumulates in proportion to the cross-sectional area of the axon, which scales with the square of the radius.

Another noteworthy computational work on single axons that extends on the above is (Barach et al., 1985), which demonstrates that simultaneous measurement of the transmembrane potential and magnetic field allows one to use the core-conductor (or “cable”) model to quantify cable constants such as axial resistances and membrane capacitance. These constants could then be used to predict the electric or magnetic field generated by an axon when only one is known.

**SQUID / MEG based experiments**

Roth and Wikswo described the first experiment in which the transmembrane potential and magnetic field of a crayfish medial giant axon were simultaneously measured (Roth and Wikswo, 1985). Prior to this work, experiments to measure each were done separately, and the results were compared using theoretical predictions. Here, the authors verified with experimental data past theoretical findings from (Woosley et al., 1985) that showed that the magnetic field can be predicted from the transmembrane potential and vice versa. To measure the potential, two electrodes were placed: one along the central axis of the axon and another approximately 30 mm from the axis inside a Tris buffer bath. Magnetic measurements were taken by threading the axon through a ferrite toroid and recording the induced potential due to the magnetic flux passing through the toroid in a coil wrapped around it. (The properties of this toroidal ferrite magnetometer are discussed in great detail in (Gielen et al., 1986). In particular, they showed that for a crayfish axon with a 100 µm diameter, a signal-to-noise ratio of 10:1 can be expected.) Theoretical calculations and experi-
mental measurements differed by a maximum of 17% for the peak-to-peak potential amplitude and by as little as 3% for the asymmetry in the shape of the magnetic signal. Errors could be attributed to parameters in the volume conductor model which was used for the theoretical calculations, toroid effects (since the potential induced is not necessarily that at a single point along the axis of the axon, while the theoretical calculations are), inhomogeneities in the external medium which result in boundary conditions that were not accounted for, and deviations of the axon from what is expected by theory. The work provided a starting point for experiments involving simultaneous measurements that could help solve the forward and inverse problems involving electric potentials and magnetic fields.

(Van Egeraat et al., 1993) reported on the magnetic field of a crayfish medial giant axon that has been subjected to a nerve crush injury. They compared their measurements with those of a theoretical model based on Hodgkin-Huxley equations. This was critical since previous results on injured neurons only consider models for which the intra- and extracellular ion concentrations are constant. Results showed that resting potential is elevated while the action current and potential amplitudes are reduced. In addition, the shape of the action current changes from biphasic to monophasic when it approaches the site of the injury. Based on their model, they predicted that the crush seals itself with a time constant of about 45 sec. The current density is also reduced to about 0.1 mA/mm².

For a more detailed review of single axon studies, see (Wijesinghe, 2010, 2014).

2.3.2 Nerves

We next consider nerves, which consist of a bundle of myelinated axons. Each axon is wrapped in a membrane known as the endoneurium. In the peripheral nervous system, groups of axons form fascicles (surrounded by the perineurium), and groups of fascicles from a nerve (surrounded by the epineurium). In the central nervous
system, myelinated axons form bundles known as tracts.

Computational

By the late 1980s, interest had developed in modeling not just magnetic fields of individual neurons but also of nerve bundles. Stegman et al. had reported on the magnetic field due to an axon in the center of an unsheathed bundle, but the field due to axons off center had not yet been studied (Stegeman et al., 1979) or had grossly simplifying assumptions about the shape of the action potential traveling along the axon (Clark and Plonsey, 1968).

Roth and Wikswo utilized the volume conductor model to study the effects of neighboring axons on the magnetic field induced by an action potential traveling along a single axon in a bundle surrounded by a myelin sheath (Roth and Wikswo, 1985). They considered an infinitely long cylindrical axon and assumed that the shape of the action potential was completely independent of the eleven parameters characterizing the biophysical properties of the nerve bundle. Each component of the model – namely, the axon, the rest of the bundle, the sheath, and the bath – had a separate parameter for conductivity. Moreover, the bundle was treated as anisotropic, with differing radial and axial conductivities.

As in (Woosley et al., 1985), it was found that the magnetic field is directly proportional to the conductivity of the axon. Isolating the effect of the sheath by treating the axon, bundle, and bath conductivities as isotropic and identical, the authors found that when the resistance of the sheath is much less than that of the bath, the sheath has little effect on the field. When the sheath wraps tightly around an axon, it can be treated as a layer of Schwann cells, resulting in a magnetic field about the same as that of an unmyelinated axon. On the other hand, when the sheath is large, it could act as the peri- or epineurium, decreasing the amplitude by even an order of magnitude. As the internal resistance, the resistance of the sheath,
and the axial resistance of the bath start to converge, the net magnetic field decreases due to weaker current flow. Shifting the axon off-center results in very little change to the magnetic field unless the axon is placed close to the edge, in which case the fields increase (as if the sheath were not present).

Next, the authors made the bundle conductivity anisotropic to incorporate the effects of the bundle on the field. They found that effects are very similar to those of just the sheath, and depending on how anisotropic the bundle is, it can have a profound effect on the field induced by the axon ($\sigma_z$ large) or almost none at all ($\sigma_z$ small). Furthermore, the closer the axon is to the edge of the bundle, the stronger its field, and vice versa. (Note that the conductivity parameters of the bundle are macroscopic and were educated guesses. At the time, no calculations had been performed to determine the conductivity from a bottom-up approach; thus, the authors relied on an estimate made by Stegeman et al. (Stegeman et al., 1979). To date, no studies have been conducted to calculate / measure these properties.)

Wiksw o et al. (Wiksw o and Roth, 1988) bridges the gap between magnetic fields induced by a single nerve and the measurement of magnetic fields at a more superficial level in order to study whole-brain magnetic signals. Specifically, the authors modeled how the distance of the sensor from a neuron affects the ability to back-calculate the electric potential distribution across the neuron. They find that when the distance is large compared to the diameter of the axon, the electric potential is almost impossible to predict. Their modeling approach takes a similar one to those above, in which the magnetic field is a convolution of the intracellular current with a weighting function that is medium- and fiber-to-sensor-distance specific. Note that here, they considered graded potentials rather than action potentials, and the fields arise from closely packed dendrites rather than from axons. Their method showed that a SQUID with a 20-30 mm pick-up coil diameter placed 20-30 mm from the cortex could provide sensitivities on the order of fT/$\sqrt{\text{Hz}}$ and would be able to detect
populations of dendrites on the order of $10^3$. The limiting factor in the calculation of the electric potentials in the brain from these ultra-precise measurements is the high degree of anisotropy and inhomogeneity in the brain, which makes modeling the tissue highly challenging and computationally intensive. Such a study has not been attempted to date.

**SQUID / MEG based experiments**

*Experimental advances*  Wikswo et al. measured the magnetic field from an isolated frog sciatic axon using a SQUID magnetometer at room temperature (Wikswo et al., 1980; Barach et al., 1980). The diameter of the axon was about 0.4 mm and contained approximately 2000 individual nerve fibers inside a sheath. The field at 1.3 mm was found to be on the order of 100 pT and was measured with a signal-to-noise ratio (SNR) of 40:1. Peak currents were on the order of 10 $\mu$A. The high SNR was achieved by utilizing a small pickup coil and averaging of 1024 independent measurements. This was the first demonstration of measuring magnetic fields on an axon from a propagating action potential. To perform the measurement, the axon was excised from the frog and part of it was immersed in a saline Ringer bath. This bath was grounded to prevent capacitive coupling effects from contaminating the signal, which was an issue in previously published attempts. Stimulation was performed by a pair of electrodes which were placed along a part of the axon still in air. Inside the bath, a toroidal ferrite core surrounded a small segment of the axon. A pickup coil was wound around the core, and the induced voltage was fed to a SQUID magnetometer with sensitivity on the order of $10 \text{ fT}/\sqrt{\text{Hz}}$. An important consideration in the design of the pickup coil was the fact that the SQUID is a current-sensing rather than voltage-sensing device; while having fewer windings would minimize the current, too few windings would deplete the SNR.

Soon after, the authors reported on a device which allowed for measurements to be
Peripheral nerve regeneration is the process by which a peripheral nerve undergoes repair and healing after an injury. The magnetic signature of peripheral nerves after injury and during regeneration have important clinical significance for diagnostic purposes. The first noninvasive localization of a peripheral nerve in humans was performed by Trahms et al. (Ernê et al., 1988; Trahms et al., 1989). Experiments were performed in a magnetically shielded room. Healthy patients were stimulated with a 10 Hz 15 mA pulse at a distal location on their median nerve. A SQUID was used to measure the magnetic signal resulting from the stimulation, and a simple five-parameter model which assumes the source of the magnetic field is a traveling current dipole was used to back-calculate the current source from the field measurement. Individual traces had magnetic field strengths of about 100 fT; averaging of thousands of measurements was performed to increase SNR. From the measurements, the authors were able to localize the nerve in the tissue and were able to confirm their predictions with x-ray computer tomography (deviation of less than 2 mm).

Kuypers et al. reported in two studies on the magnetic field of neurons undergoing the regeneration process. This is critical in clinical settings in which the success rate of a second nerve reconstruction is significantly lower when poor regeneration is detected early. They were first interested in developing quantitative techniques
to determine how well an axon had regenerated post-injury (Kuypers et al., 1993). Existing methods relying on electrical signals were inaccurate and not capable of detecting the status of the regeneration process early enough due to variations in tissue property and geometry and the distance between the electrode and nerve, but magnetic signals tested during their experiments had a much higher fidelity, with results matching those of a histological study. Two invasive methods involving measuring the nerve compound action potentials using a ring electrode or measuring the compound action current using a magnetic sensor were tested on transected and reconstructed rabbit sciatic nerves. Amplitudes of the measured signal were found to be more consistent and stimulus artifacts were smaller in the magnetic recordings compared to electrical. Regeneration times were much more predictable than with electrical recordings because the impedance of the tissues causes signal variation and/or attenuation.

Kuypers et al. also showed that magnetic measurements of neuronal activity provide complementary information about remaining functional neurons to that provided by histological studies (Kuypers et al., 1998). Utilizing the measurement technique developed in (Wikswo et al., 1980, 1991), the authors quantified the number of neurons in a rabbit peroneal nerve still functioning after transection and reconstruction. This involved measuring compound action currents by threading the nerve through a toroidal ferrite core which was wound with a pickup coil. After recording, the animals were euthanized, and a histology was performed to count the number of remaining cells. An overall reduction of 5% of cells and 50% of magnetic field compared to baseline suggested that about 50% of the neurons in the nerve had either lost function or were completely degenerated as a result of the transection.

Fukuoka et al. performed an in vitro study in which the dynamics of the compound action fields and compound action potentials were examined near the site of a nerve block (Fukuoka et al., 2002). The study was performed on rabbit sciatic
nerves, which were electrically stimulated, and a current dipole model was used to determine the source of the current at several locations along the nerve and at several time points. The peak of the field which was present in healthy parts of the nerve was absent at the site of the injury. A related study was performed by Hoshino et al., who were able to perform three-dimensional localization of a lesion to the cervical spinal cord of rabbits via a 24-channel SQUID (Hoshino et al., 2005). This was done by applying electrical stimulation and measuring evoked magnetic fields at 64 to 96 points along the cervical spine. More recently, (Mackert, 2004) has reviewed several successful endeavors and promising avenues for 3D magnetic localization of nerves and nerve-related pathology including high-resolution mapping of the brachial plexus with SQUIDs and localizing conduction blocks (which could be a symptom of nerve injury). Methods involve similar techniques to those discussed above such as measuring the compound action current, developing a current source model, and reconstructing the source using estimates of the current, conduction velocity, and other relevant properties of the tissue.

For a more detailed review of nerve studies, see (Wijesinghe, 2010, 2014).

**Clinical applications** Clinical problems offer a great source of motivation for scientists studying nerves and nerve bundles because of the potential for direct impact on patients’ lives. Wikswo et al. described the challenges involved in making in vivo measurements of human nerve function during surgical procedures, including challenges in placing electrodes, uncertainty in extracellular conductivity, and the risks of elevating the nerve into air for recording (Wikswo et al., 1989). In their work, they developed a probe that can be placed around a nerve to measure activity of the axons within it. Patients were being treated for carpal tunnel syndrome, a debilitating condition of the hand and wrist resulting from compression of the median nerve (Office of Communications and Public Liaison, 2018). Only a short segment
of about 2 cm of the nerve was exposed to air while the remainder was immersed in physiological saline. Electrical stimulation at an amplitude of 2 mA and 100 µs pulse width was administered, and measurements were bandpass filtered to between 1 Hz and 10 kHz. Repeated measurements were taken to eliminate noise and artifacts. Measured compound action potentials had an associated current of about 650 nA. The measured conduction velocity of one patient of 55 m/s suggests - which falls on the lower end of the normal range of 55 – 73 m/s - suggests that age, temperature, interrupted circulation, or disease processes may have affected the nerve.

Localization of conduction blocks is an important application of MEG that has received significant attention over the past two decades. (Mackert et al., 1998) reported on the noninvasive identification of conduction blocks in the tibial nerve through evoked potentials. A multichannel SQUID consisting of 63 channels was used to image patients in the lower lumber and upper sacral areas of the spinal cord. Patients had suffered various forms of acute trauma that resulted in loss of reflex and/or muscle control. A dipolar source model was used to calculate the flow of compound action potentials, and results showed either slowing or block of conduction on the side of the spinal cord affected by injury. The imaging was performed in 15 minutes, a reasonable amount of time in which to diagnose a patient in the clinic. Similarly, Ishii et al. reported on measuring magnetic signals due to stimulation of the tibial nerve in healthy male subjects (Ishii et al., 2012). Unique to their study was the fact that, for validation, percutaneous electrodes were placed in the lumbar epidural space to measure the electrical activity there as a result of stimulation; measurements of electrical signals corroborated results of the magnetic measurements.

Extending work they had reported in (Fukuoka et al., 2002), Fukuoka et al. studied rabbit sciatic nerves with incomplete conduction blocks as this is a more common clinical case than complete conduction block and can be more difficult to
identify (Fukuoka et al., 2004). A similar preparation procedure in which sciatic nerves were excised, placed in saline Ringer solution, and stimulated to evoke compound action potentials and compound action fields was performed. A significant reduction in their amplitude and conduction velocity was noted in comparison to before the injury. The method allowed detection of the injury to within a 3 mm resolution. Incomplete blocks directly in the spinal cord had not been studied until the report by Tomori et al. (Tomori et al., 2010). This is common in conditions such as myelopathy, which is excessive compression that occurs in the spinal cord due to trauma, congenitive stenosis (narrowing of the spinal cord), or disc degeneration / herniation. To perform their experiment, a balloon catheter was inserted epidurally in the cervical spine of rabbits. Evoked potentials and magnetic fields were recorded with epidural electrodes and a SQUID magnetometer, respectively. Similar to the results reported by Fukuoka et al. for peripheral nerves (Fukuoka et al., 2004), a marked decrease in amplitude and conduction velocity at the site of compression was identified.

2.3.3 Whole brain and spinal cord

Finally, we consider studies of the whole brain and spinal cord. These are typically performed with the goal of mapping sensory projections and identifying sites of lesions and injury.

Computational

MEG (Blagoev et al., 2007) modeled the magnetic field of anatomically realistic, physiologically idealized neurons and found that the field far from the neurons due to stimulus-induced activity can be well-approximated by a dipole field while inside the cortex, the approximation was poor. They employed the software BioSENSE for their simulations, which was capable of modeling dynamics in 3D compartmentalized
neurons. (Zumer et al., 2008) developed a probabilistic method to localize neural activity via MEG (and electroencephalography). Unique to their approach was a set of temporal basis functions, which are learned from the data via factor analysis. These basis functions were combined to give an estimate of neural activity in each voxel. To automate the process of whole brain cortical activity reconstruction, (Dalal et al., 2011) developed a toolbox known as NUTMEG which provides access to different reconstruction algorithms and a graphical user interface to make analysis more user-friendly.

**nc-MRI** Similar to approaches taken by researchers studying MEG, (Konn et al., 2003) modeled the brain as a conducting sphere and the currents that flow within it as a simple dipole. This dipole produces a spatially varying magnetic field whose phase varies throughout the sphere. The values predicted by their simulation for minimum dipole strength (4.5 nA-m) and minimum detectable field strength (100 pT) matched those expected by spontaneous neural activity. (Xue et al., 2006) extended on this work by improving the dipole model. In their approach, each dendrite, which is where the majority of current is sourced in a myelinated axon (Swinney and Wikswo Jr, 1980), can be modeled as an individual dipole system, two dipole system (in which the interactions between dendrites are considered), or multi-dipole (in which all the dendrites within a given voxel of an MRI are considered together). They accounted for spatial differences in the way the magnetic field is calculated by composing a piecewise function for magnetic field strength based on distance from the source. Their results showed that unlike MEG, nc-MRI is insensitive to phase delays between the activities of different neurons. It is, however, sensitive to other factors like dendritic parameters such as number, size, and the number firing during the echo time of the MRI (the time during which signal is acquired) and the orientations of the neurons relative to $B_0$. (Heller et al., 2009) offered additional
support for nc-MRI by developing an analytical model, comparing it to simulations, and showing that their analytical model agrees with simulations.

The utility of nc-MRI has been heavily debated (Hagberg et al., 2006), and other scientists offer conflicting viewpoints. While (Xue et al., 2006) claimed that their dipole model also applies to unmyelinated axons, (Park and Lee, 2007) showed that the intracellular current in an axon follows a quadrupole configuration. This in turn results in a bipolar magnetic field, and due to cancellations, no change in the MRI is to be expected due to axonal current. This result, however, disagrees with the experimental evidence produced by (Roth and Wikswo, 1985), who showed that axonal currents produce magnetic fields on the order of 100 pT. On the other hand, the results agree with those of (Xue et al., 2006) for dendrites since the dendritic magnetic field is unipolar, resulting in summations. (Park and Lee, 2007) also suggested that unless the nc-MRI signal can be separated from the BOLD signal evoked by fMRI, the signal from nc-MRI would get washed out. (Jay et al., 2012) estimate the phase shift due to dendritic activity, and show that only when tens of thousands of dendrites are firing simultaneously can the signal be detected by conventional MRI. Thus, it is unrealistic to expect nc-MRI to detect neuronal magnetic fields given the present state of MRI. Similar to (Blagoev et al., 2007), (Luo et al., 2011) developed an anatomically realistic model based on human neurons, simulating dendritic branching of pyramidal neurons in the cerebral cortex and their physiological properties. The authors found that evoked potentials are too difficult to detect while spontaneous activity may be possible to detect but only under optimal experimental conditions.

SQUID / MEG based experiments

Landmark discoveries
Magnetoencephalography: Detection of the Brain’s Electrical Activity with a Superconducting Magnetometer (Cohen, 1972) Cohen used a SQUID magnetometer to measure magnetic fields generated by neural activity in the brain on a single-trial basis, coining the term magnetoencephalography. Previous experiments to measure the magnetic field of the brain had relied on averages performed over several trials using coils that were placed inside a magnetically shielded room. Cohen showed that magnetic fields provide information that is unique from that provided by EEG because they are capable of detecting direct current (DC) and distinguishing between situations in which EEG is nonzero while the currents are still flowing and vice versa. Cohen confirmed that MEG signals are temporally correlated with EEG during eyes closed and open states. The strength of the magnetic field about 5 cm above the scalp was on the order of 1 pT.

Visually evoked magnetic fields of the human brain (Brenner et al., 1975) Magnetic field measurements from a SQUID were used to determine the cortical response to the presentation of vertical bar stimuli at a 10 Hz frequency. This was the first paper to describe evoked magnetic responses to any kind of stimulus. The SQUID was most sensitive to stimuli in the range of 5 to 25 Hz. The experiments demonstrated that because magnetic fields can penetrate through the skull with minimal attenuation due to currents flowing in the dermis, they provide a more spatially measurement of cortical neuronal activity compared to measurements of electric potential. The authors found that the visually evoked response is about 10 times weaker than that of spontaneous neural activity, falling in the 0.1 pT and 1 pT ranges respectively.

Somatically evoked magnetic fields of the human brain (Brenner et al., 1978) Magnetic fields that were triggered as a result of electrical stimulation applied to the fingers of four human subjects were reported. The authors found that the field is
localized to areas in the brain which receive projections from the skin and has a magnitude on the order of pT within millimeters from the scalp. This is in contrast to the visual evoked potential (VEP) which are diffuse signals that are spread across the scalp. The magnetic fields were measured with a SQUID magnetometer. Frequencies of stimulation between 3 and 30 Hz were tested; for frequencies above 18 Hz, the phase response of the magnetic fields was linear, whereas for frequencies below 18 Hz, the frequency response was nonlinear and exhibited considerable variability among the subjects. The latency of the magnetic field with respect to the stimulus was about 70 ms. This together with a reaction time of about 172 ms for the same cohort results in a motor response time of about 102 ms, which closely matches results reported in previous literature of 115 ms.

Human magnetic auditory evoked fields (Reite et al., 1978) Reite et al. presented measurements of evoked magnetic fields due to auditory stimuli. Four subjects were recruited and an asymmetric second derivative SQUID gradiometer was used to measure the evoked potential in temporal and central areas of the brain (T3/T4 and C3/C4 in EEG coordinates). Using a headset, the subjects listened to clicks which were generated by 1 ms square wave pulses. Recordings were highly sensitive to location, and the signal to noise ratio greatly fell if the coil was moved either 4 cm anterior or posterior to these locations; this is in contrast to EEG recordings which were distributed throughout the scalp. Averaged evoked fields were localized and well-defined but differed across subjects. Field magnitudes at a recording distance of 2 cm above the scalp were about 1 pT and appeared at a latency of about 45 ms.

Experimental advances Since the landmark experiments described in the previous section, a number of investigations have also been performed on using MEG to characterize brain activity. (Yang et al., 1993) performed a study to demonstrate the re-
liability and anatomical validity of somatosensory source localization via 37-channel MEG. Experiments were performed on two male subjects on a 144 mm diameter circular area over the parietotemporal cortex. Tactile stimulation was performed on the contralateral side on 66 locations on the face, hand, and arm with the goal of mapping the target of somatosensory stimulation in the brain. The localization of the source of the magnetic signal matched closely with expectations based on an anatomical MRI that was also performed on the subjects. (Kakigi et al., 2000) reviews a number of related contributions to somatosensory mapping.

(Hämäläinen and Hari, 2002) reviews many studies that report on measuring the dynamics of neural activity via MEG. EEG and MEG are the only two mechanisms that are capable of providing sub-millisecond temporal resolution for neural recordings. fMRI fails here since the hemodynamic effect is typically measured with a time resolution of seconds. As has been described above, MEG offers an advantage over EEG since the scalp can distort EEG signals while magnetic signals can be measured without the scalp as a source of disturbance. Studies that were reviewed include those characterizing a sequence of evoked responses, cortex-muscle coherence, dynamics of mirror neurons, binaural hearing, and preoperative functional localization.

Clinical applications Advances in MEG devices and protocols have a number of potential clinical use cases including conditions such as presurgical mapping of tumors and diagnosis of epilepsy, stroke, and nerve injury (Wikswó, 1995).

Imaging of the dynamics in the healthy spinal cord, known as magnetoneurography (MNG), is an active area of clinical research. (Mackert et al., 2000) reported three-dimensional mapping of the nerves of the human brachial plexus in healthy patients using a 49-channel SQUID magnetometer. Propagation of compound action currents was used to distinguish individual nerves in the plexus and measured conduction velocities around 56 m/s. Recently, a 120-channel SQUID was used to
measure evoked magnetic fields in the cervical spinal cord due to stimulation of nerves in the thoracic spinal cord (Sumiya et al., 2017). Healthy subjects showed no decrease in conduction velocity while subjects with lesions showed decreased conduction velocity. The site of stenosis in a patient with cervical spondylotic myelopathy was mapped by the MNG.

While lesions in the central and peripheral nervous system are a focal point for MEG diagnostics, MEG is also a useful tool for predicting seizures. (Mikuni et al., 1997) reported on using whole-head MEG and subdural electrodes to simultaneously record epileptic events in two patients with temporal lobe epilepsy. In one patient, MEG and electrode recordings matched closely and the MEG was able to localize the tumor with high accuracy. Moreover, the greater the amplitude of the epileptic spikes and the more lateral their origin, the greater their detectability.

MEG can also be used to diagnose Alzheimer’s disease (AD), which affects 1.7% of the American population (Alzheimer’s Association, 2018). Fernandez et al. designed a functional biomarker for AD in which the power of different bands of an MEG are evaluated (Fernández et al., 2006). In their study, 22 patients with AD and 21 healthy subjects were examined with a 148 channel whole-head MEG during a 5-minute resting period. The spontaneous MEG was analyzed via power spectral density; the signal was bandpass filtered between 2 and 60 Hz into 2 Hz windows. A statistical analysis showed that of all possible bins, those between 16-28 Hz exhibited the greatest difference in power between healthy subjects and AD patients, resulting in a sensitivity of 81% and specificity of 80%. These results suggested that the beta band signal is a potential biomarker for AD. A result by Stam et al. showed a similar effect; they took a graph theoretical approach in which phase lags were computed between MEG recordings taken on different channels in a whole-head MEG (Stam et al., 2008). AD patients showed a lower phase lag index in the lower alpha and beta bands compared to controls. The mean clustering coefficient and path length -
measures of how similar nodes are to each other - were also significantly smaller in patients with AD. A number of other MEG based biomarkers are reviewed in (Laske et al., 2015).

*nc-MRI based experiments*

Despite the controversy in the feasibility of nc-MRI, various authors have attempted to use it to map neuronal activity. The original demonstration of nc-MRI was performed by (Joy et al., 1989). The authors showed both in vitro in a phantom and on a human subject that by measuring the phase of protons that are initially precessing in synchrony but are deflected due to an RF pulse, one can generate an image in which the magnitude of the signal at each pixel is proportional to the phase shift. This is the fundamental operating principle of nc-MRI. They showed that in vitro currents can be detected with a spatial resolution of 1 mm and a precision of 2.5 µA, and electrical stimulation with 2 mA current can be detected in vivo.

A number of variants of this basic principle have been proposed over the years to optimize the quality of the signal acquired. (Bodurka and Bandettini, 2002) later show that by utilizing spin-echo (SE) echo-planar imaging (EPI), which dephases protons differently than the standard method described above, one can detect transients as short as 40 ms and magnetic fields as small as 200 pT. Measuring phase shift has been a common approach employed by scientists using nc-MRI to detect alpha waves in subjects whose eyes were closed (Konn et al., 2004) and to image peripheral nerves in nonhuman subjects (Wijesinghe and Roth, 2009). (Truong and Song, 2006) also applied oscillating gradients in an MRI scan to detect a loss of phase coherence due to differential Lorentz forces acting on protons. They found that the procedure greatly enhances sensitivity and were able to achieve millisecond temporal resolution during measurement of the human median nerve.

(Xiong et al., 2003) directly compared their measurements to that of fMRI and
show that their measurements are consistent but offer a much higher temporal and spatial resolution (100 ms and 3 mm, respectively). Moreover, they showed that the origin of neural activity as well as temporal latency for each of their visual, motor, and somatosensory experiments closely matched their expectations based on task design and reports in literature. Essential to their approach was their decision to measure magnitude rather than phase coherence of the RF magnetic signal generated by the precessing protons. Future studies, including those by Chow to measure signals in the human optic nerve, successfully followed this precedent of measuring magnitude (Chow et al., 2006a,b).

Three recent studies also reported on the infeasibility of nc-MRI due to low contrast in comparison to the BOLD background signal generated by fMRI. (Parkes et al., 2007) took a two-pronged approach to verifying the utility of nc-MRI. They first aimed to replicate the prominent visuomotor experiment performed by (Xiong et al., 2003) described above. Experimental conditions were closely replicated except for a minor difference in the interstimulus time (approximately 1 second on average instead of 2) which minimizes BOLD signal and increased jitter which helps subjects pay attention during the task. Second, they aimed to detect steady state visual evoked potentials (ssVEPs) via nc-MRI. Their results showed that in neither experiment was the MRI machine sensitive enough to distinguish the magnetic source signal from the BOLD signal. In another experiment by (Huang, 2014), visual stimulation was provided to a monkey whose neuronal firing patterns for the stimulus are known. However, no magnetic source signal could be observed. The same was true for a similar VEP experiment performed in an octopus, which does not have hemoglobin-containing blood, despite using a powerful 9.4 T scanner (Jiang et al., 2014).

Thus, both computational and experimental evidence provide conflicting opinions as to the efficacy of nc-MRI as a means of directly imaging neuronal activity. Although magnetic signals certainly are generated by neurons and the theoretical
benefits of nc-MRI including high spatial and temporal resolution would be nice to have, more studies must be conducted before the technique can be accepted as a valid experimental tool.

2.3.4 Promising new magnetic sensing / imaging technology

While SQUID has for decades offered unparalleled precision in its ability to measure weak magnetic fields (with a sensitivity of up to $5 \text{ fT}/\sqrt{\text{Hz}}$ for a bandwidth of 100 Hz) (Kominis et al., 2003), two very recent technologies known as atomic magnetometers and nitrogen-vacancy (NV) defect center-based magnetometers are showing great promise in replacing SQUID as the new gold-standard for biomagnetic imaging. Interestingly, both methods allow sensing to be performed at room temperature, opening doors for many possible commercial applications.

Atomic magnetometers

Atomic magnetometers operate based on the precession of spin-polarized atoms – typically an alkali-metal vapor – in a magnetic field. These devices have sensitivities on the order of $0.3 \text{ fT}/\sqrt{\text{Hz}}$ for a bandwidth of about 1 Hz. (Kominis et al., 2003) described a variant called a spin-exchange relaxation-free (SERF) magnetometer which operates using high alkali-metal density (with pressures reaching several atmospheres) and low magnetic field. They demonstrated experimentally that the device has a sensitivity of $1 \text{ fT}/\sqrt{\text{Hz}}$ in the range of 10 to 150 Hz and $0.54 \text{ fT}/\sqrt{\text{Hz}}$ in the range of 28 to 45 Hz. The authors claim that non-invasive studies of 0.1 to 0.2 mm size “modules” of the brain could be performed after some optimization.

Development of atomic magnetometers is still underway, but researchers have already demonstrated its utility and adaptability for many applications. (Schwindt et al., 2004) developed a chip-scale atomic magnetometer with a sensitivity of 50 pT/$\sqrt{\text{Hz}}$ at 10 Hz. While this is two orders of magnitude lower than the state of the
art SERF magnetometers, the device is only 12 mm$^3$ in volume and consumes only 195 mW of power. The device can be battery operated and manufactured at low cost. Taking this idea one step further, (Sander et al., 2012) developed a 1000 mm$^3$ atomic magnetometer to measure the brain’s magnetic fields. Because the chip is much larger than that of Schwindt et al., it can be easily manipulated by hand and used in a similar way to an EEG electrode. Both spontaneous and somatosensory evoked responses were measured, and the authors found that the measurements compared closely with those made by a SQUID. The device works by measuring the absorption of infrared light by rubidium gas enclosed in a 0.77 mm$^3$ chamber. The sensitivity of the device is on the order of 200 fT/$\sqrt{\text{Hz}}$ compared to 0.2 fT/$\sqrt{\text{Hz}}$ for SQUID sensors. However, the device is still capable of measuring pT magnetic fields, on the order of the magnetic fields produced by neural activity. Finally, (Kim et al., 2014) reported a 256-channel atomic magnetometer with a sensitivity of 4 fT/$\sqrt{\text{Hz}}$ to measure auditory evoked fields.

*Nitrogen-vacancy (NV) defect center-based sensors*

Nitrogen-vacancy (NV) defect center-based sensors are a counterpart to atomic magnetometry, offering incredible spatial resolution at the expense of magnetic field sensitivity. These devices operate by coherent manipulation via electron spin resonance of qubits. Optical pumping allows for polarization of a nitrogen vacancy in diamond; this polarization can then be measured through a technique called state-selective fluorescence. (Maze et al., 2008) showed that they are able to detect magnetic fields of up to 3 nT at kHz frequencies; moreover, they found a sensitivity of 0.5 $\mu$T/$\sqrt{\text{Hz}}$ with a crystal of size only 30 nm (the resolution of the device). The device is photon shot noise-limited, and the longer the averaging time, the lower the magnetic field strength it is able to measure, with a sensitivity of about 2 nT after 100 sec of averaging. Similar results were reported by (Balasubramanian et al., 2008). Since the
magnetic field of a single electron 10 nm away produces a magnetic field of about 1 \( \mu \)T, it may be possible to detect these spins with such a device. Furthermore, as suggested by (Taylor et al., 2008), at a distance of 10 nm, the field from a single proton whose magnitude is on the order of 3 nT should be measurable if certain modifications such as increasing the density of nitrogen vacancies, introducing bound electron substitution impurities, and employing advanced forms of dynamical decoupling are made.

(Dolde et al., 2011) reported on a NV defect center-based sensor capable of measuring electric fields with a sensitivity of \( 202 \pm 6 \text{ V}/(\text{cm} \cdot \sqrt{\text{Hz}}) \), enough to measure the electric field produced by an elementary charge located 150 nm from the sensor. This has major implications for neuroscience since it could allow researchers to measure the dynamics of individual ion channels with extremely high precision. To this end, (Barry et al., 2016) designed a biocompatible NV defect center-based sensors to measure action potentials of individual neurons over long periods of time in vivo. This allowed them to see how action potentials propagate at a submillisecond time resolution and to monitor individual neurons without performing genetic modifications.

2.4 Non-synaptic interaction among neurons

While synaptic communication predominates in the nervous system, evidence also suggests that neurons can communicate non-synaptically via electric fields. Furthermore, theories of the role of magnetic fields in neuronal computation and memory have also been proposed. We discuss these topics here.

2.4.1 Electric

The term ephapse was first used by (Arvanitaki, 1942) to describe the effects of cortical electric fields on the membrane potentials of neurons. These electric fields
exist due to the presence of ions and fluctuate over space and time as the distribution of ions changes. Recent studies have focused on the idea of ephaptic coupling, which suggests that extracellular electric fields can synchronize the activity of a group of neurons (Anastassiou et al., 2010, 2011).

To investigate the phenomenon, they modeled the effect of spatially inhomogeneous electric fields and apply the model to study neurons in the rat hippocampus (Anastassiou et al., 2010). According to the study, the transmembrane potential has a very complex relationship with the extracellular field depending on the spatiotemporal structure of the field. While for a spatially homogenous field, the transmembrane potential is impacted only due to the source/sink boundary conditions of the axon, for an inhomogeneous field whose spatial frequency varies more rapidly than the length of the axon, the transmembrane potential depends on the characteristic length of the external field, the cable length, and the electrotonic length (which is the relative size of the cable to the space constant).

The authors investigated the phenomenon of ephaptic coupling and made several conclusions about the effect of extracellular fields on spike rate and timing. Action potentials are a significant source of the extracellular field since a large number of ions are exchanged between the cell and the surrounding medium. However, according to the study, extracellular fields tend to have almost no effect on the membrane during an action potential. Moreover, even if two neighboring neurons show coherent oscillations in their membrane potentials in the presence of an electric field, their spike times may be shifted. One notable exception to this is during epileptic seizures, during which the synchronized activity of many neurons can activate neighboring subthreshold neurons. Typically, fields are not strong enough to take a neuron from rest to firing, while at the same time, once a neuron is in the process of firing, fields have little to no effect.

(Anastassiou et al., 2010) stressed the importance of LFPs in synchronizing pools
of neuronal activity. In the context of the rat hippocampus, while theta (6-10 Hz) waves tend to have little effect, sharp waves, which have high temporal frequency between 140-200 Hz, tend to couple pyramidal neurons. The shift in spike phase could be as large as 36 to 72 degrees and is theorized to occur at the dendrite, with the effect propagating to the soma and remaining part of the neuron. Thus, theta waves are like fine tuners while sharp waves are more akin to a synchronizing mechanism. These results suggest that LFPs can affect spike timing via ephaptic coupling and highlight the two-way nature of induced electric fields in the brain, in which the neurons generate fields and fields feed back onto neurons.

In another set of experiments to study the properties of ephaptic coupling, a stimulating electrode was placed 50-150 micrometers from a neuron cell body; intracellular and extracellular voltages from multiple locations were recorded to accurately characterize entrainment and the magnitude of the electric field (Anastassiou et al., 2011). Both subthreshold and superthreshold entrainments were observed. For AC current stimulation with amplitudes between 25-200 nA and a frequency of 1 Hz, electric fields ranging from $0.74 \pm 0.53$ to $5.86 \pm 4.25$ mV/mm and membrane potentials between $0.16 \pm 0.005$ to $0.14 \pm 0.007$ mV were observed. These magnitudes were constant over frequencies up to 100 Hz. Furthermore, the measured membrane potential was approximately 180 degrees out of phase with the extracellular potential, a relationship which persisted even when a subthreshold intracellular current was injected into the cell. When the neuron was forced into a superthreshold state by current injection, a comparison of spiking activity before and after AC stimulation showed that spike timing was phase shifted by 270 degrees with respect to the stimulation signal. This corresponded to the rising edge of the AC signal before reaching peak amplitude. The average spike count, however, was not affected. The experiment was repeated with recordings taken from multiple neurons, and all neurons’ membrane potentials aligned 270 degrees out of phase with the AC stimulation signal. These
results suggest that synaptic input dominates over ephaptic input in causing neurons to spike. Nevertheless, even very weak electric fields due to ephaptic effects (as low as 1 mV/mm) can cause phase shifts in spike timing and induce synchrony.

2.4.2 Magnetic

(Hagberg et al., 2006) in their description of nc-MRI briefly touched on the spatial extent of the magnetic field generated by neuronal activity. Assuming the time course of neuronal currents is between 1 and 200 ms, they showed that the electromagnetic wave that propagates due to action potentials and local field potentials, respectively, has a characteristic distance $1/(t \cdot c)$ of between 10 nm and 3 µm. This has implications for nc-MRI but also for biophysical processes such as magnetic coupling of neurons. In the context of nc-MRI, variations in neuronal activity be considered more or less stationary. Considering a rectangular wave approximation to an action potential with the above characteristic distance, then for magnetic coupling, the calculation implies that neurons are only capable of synchronizing in phase with others that are within this distance.

However, (Banaclocha, 2007; Martinez-Banaclocha, 2017) have formulated an extensive theory in which magnetic fields generated by neuronal activity and glial cells are responsible for aspects of cognition such memory and computation. Moreover, they have provided a neuromagnetic explanation for the presence and intensity of spontaneous neuronal activity, which, as has been described in sections on MEG above, is often stronger than that of evoked responses to stimulation (Banaclocha and Banaclocha, 2010; Martinez-Banaclocha, 2017). While the theory remains controversial due to the lack of experimental data supporting these findings and potential misinterpretation of past results by the authors, it is nevertheless an important theory which we review here. Key to Banaclocha’s hypothesis is the so-called “cytoarchitectonic structure of the astroglial network in the human neocortex” (Banaclocha,
According to this theory, the columnar arrangement of astroglial cells and neurons in the cortex maximizes the potential for magnetic coupling between the glial cells, which supposedly maintain static magnetic fields, and neurons, which generate alternating magnetic fields. They state that glial cells, which are exponentially more prominent in proportion to neurons in humans compared to other organisms, may be implicated in complex brain functions such as memory and consciousness. This would take place through a process in which the magnetic fields generated by neurons would “write” to the static magnetic fields of the glial cells, and that these magnetic fields could be retrieved as memories. Citing the work of Roth and Wikswo, they calculated the magnetic field of an axon to be on the order of $\mu$T inside the cortex itself. This calculation agrees with ours:

$$120 \times 10^{-12} = B = \frac{\mu_0 I}{2\pi r} = \frac{\mu_0 (7.8 \times 10^{-7})}{2\pi (0.5 \times 10^{-6})} = 3.1 \times 10^{-7} = 0.3 \mu T$$

However, this authors reported this field strength at 0.5 mm while we find this field strength at 0.5 $\mu$m. The authors do not clarify exactly what magnetic field strength is needed to induce a permanent magnetic field in astrocytes, but assuming a field on the order of $\mu$T, only astrocytes located adjacent to an axon, whose diameter is assumed to be around 0.5 $\mu$m, would be affected by neuronal magnetic fields.

Another claim the authors made is that frequency of magnetic field oscillations is important for storing memories and for sensory integration (Banaclocha, 2007). Specifically, they cited past research that suggests that frequencies below 50 Hz are important for memory formation, and frequencies between 40 Hz and 80 Hz are involved in combining different aspects of a visual scene including color, shape, and movement. While the findings seem to match discoveries involving transcranial
magnetic stimulation to facilitate memory formation (Wang et al., 2014), the latter does not cite astroglial cells as playing a role in this phenomenon.

(Banaclocha and Banaclocha, 2010) also described the implications of this architecture for spontaneous neocortical activity. They cited research that reported on so-called cortical flashes, which are events involving synchronized activity that occurs in small groups of neurons in the cortex; spontaneous neuronal avalanches are similar events that occur at the LFP level. No mechanism has been agreed upon to explain this phenomenon. The authors proposed that astroglial cells undergo a self-organizing process in which spontaneous neural activity is generated as neurons fire and modulate the magnetic field of the astroglial cells which in turn affect the neuronal dynamics. This results in a chaotic circuit in which spontaneous neuronal activity is generated as the system falls into an attractor state. They then cited a theoretical discovery by Hopfield, which shows that stable states of a dynamical system can be treated as general, content-addressable memory (Hopfield, 1982, 1984). This ties back to their theory of astroglial cells being involved in memory formation and retrieval in which long-term exposure to cortical oscillations entrains astroglial cells to store information in the form of a magnetic field.

Finally, they stated the potential impact of magnetic fields in modifying the rate of influx of calcium ions into astroglial cells (Martinez-Banaclocha, 2017). Specifically, they argued somewhat incorrectly that based on studies of magnetic fields produced by currents in potassium and sodium ion transport channels that magnetic fields may be a necessary component by which these ion transport channels function. The study they referred to is (Soares et al., 2015), in which the authors developed a finite element model to account for the said magnetic fields. Soares’ model utilized the Poisson-Boltzmann equation to characterize the distribution of these electrolytes around the lipid bilayer. Then, factoring in the permeability of the bilayer and the ionic current density, they were able to write an equation in the form of Ampere’s
law for magnetic field as a function of current density in the different channel types. They found that sodium channels generate a higher field intensity than potassium channels; however, given the density of ion channels and the spatial extent of the fields generated by each type, they concluded that it is unlikely that fields induced by neighboring channels can affect the function of a given channel. Banaclocha seemed to misinterpret the words of Soares et al. who clearly state that “the existence of a number of active Na\textsuperscript{+}-channels a given membrane region [sic] does not appear to interfere directly in the functioning of K\textsuperscript{+}-channel [sic] located among them, and vice-versa, generating only a short-range perturbation” (Soares et al., 2015). Instead, Banaclocha stated, “the number of active Na\textsuperscript{+}-channels in a given membrane region produce a short-range interference in the functioning of K\textsuperscript{+}-channels, and vice versa” (Martinez-Banaclocha, 2017).

2.5 Magnetic stimulation of neuronal activity

Magnetic stimulation of the brain is known to modulate neural activity and can arise from a variety of sources. In the laboratory and clinic, transcranial magnetic stimulation, or TMS, is often employed to perform diagnostic testing or to administer a treatment. On the other hand, ambient magnetic fields such as those emanating from power lines can also modulate neural activity and morphology in strange ways. We describe these phenomena below.

2.5.1 TMS

Transcranial magnetic stimulation, or TMS, is a technique in which a large current (on the order of kA) is used to generate a magnetic field which can be focused on different parts of the nervous system to induce neural activity for scientific, diagnostic, and therapeutic purposes (Hallett, 2000). The current usually flows in a loop or butterfly shaped apparatus; a magnetic field is then generated, according to Am-
pere’s law. The shape of the apparatus dictates where the field will be focused and the strength of the field. Various TMS protocols exist, ranging from short TMS, or sTMS, in which a single pulse is delivered, to repetitive TMS, or rTMS, in which multiple pulses are delivered for a longer period of time. Each have their purposes; for instance, a short pulse can be used to measure the conduction velocity of a nerve that may be compressed while a longer pulse could be used to treat epilepsy.

(Barker et al., 1985) were the first to report on the magnetic stimulation of neurons in the human brain. The stimulator consisted of a flat coil with a 10 cm diameter and could achieve a peak current of 4 kA. Applied to the motor cortex, the stimulation could evoke a twitch which was measured by electromyography (EMG). Soon after, a number of studies followed detailing the response of hand muscles to TMS (Hess et al., 1986, 1987a), reporting on nerve conduction velocity in patients with multiple sclerosis (Barker et al., 1986; Hess et al., 1987b), and optimizing coil design and stimulation parameters (Cohen et al., 1990; Epstein et al., 1990). Below, we review a number of important scientific and clinical discoveries in TMS research with a focus on motor control and mental health, areas which have particular clinical relevance.

**Motor control**

(Hess et al., 1987a) performed one of the pioneering studies detailing hand movement in humans undergoing TMS. A magnetic field was generated by a 9 cm circular coil producing a magnetic field between 0.9 and 1.6 T. The authors reported contraction of three separate muscles in the right hand and showed that greater magnetic field strengths resulted in stronger contractions and lower latency. Voluntary background contraction resulted in reduced latency and activation of spinal motoneurons when stimuli were applied both at the brain and at the wrist. Interestingly, voluntary contraction of the ipsilateral muscles also resulted in lower thresholds and latencies.
for muscles on the contralateral side of the stimulus.

After performing a seminal study in repetitive (rTMS) to modulate speech, Pascual-Leone reported an rTMS study of the motor cortex of 14 healthy human subjects (Pascual-Leone et al., 1991). Unlike the short TMS (sTMS) employed by (Hess et al., 1987a), rTMS was able to evoke both excitatory and inhibitory effects in the hand. Moreover, the effect was less localized, activating an entire group of muscles near the target site, although the latency of activation of the neighboring muscles was larger than the corresponding latency if those sites were targeted directly. Also unlike sTMS, the greater the intensity and frequency of stimulation, the fewer the number of evoked potentials measured in the target. rTMS resulted in a period of heightened sensitivity during which sTMS was able to evoke potentials in the EMG with a greater amplitude. The authors hypothesized that repeatedly applying rTMS would result in the reduction of inhibitory activity and recruitment of excitatory pathways.

(Chen et al., 1997) demonstrated that rTMS could be used to induce long-term depression in motor neuron circuitry in humans. Previous results had shown that very low frequency stimulation (around 1 Hz) could induce depression in both cortical slices and in vivo animal studies. In this experiment 14 subjects were stimulated with a figure-eight shaped coil with two different stimulation patterns: 0.1 Hz for 1 hour and 0.9 Hz for 15 minutes. EMG activity from the contralateral wrist, biceps, and shoulder was recorded. Results showed that for the first group, motor evoked potential amplitude remained constant over the course of the session while for the second group, it decreased, implying that depression had occurred. These findings have ramifications for treating conditions involving neuronal hyperexcitability including epilepsy and myoclonus.

In a landmark study, (Huang et al., 2005) reported on the utility of rTMS to generate theta band activity that can generate long-term synaptic plasticity in the
motor cortex. Three protocols known as cTBS, iTBS, and imTBS were tested: each consisted of the same number of total pulses delivered at 50 Hz but differed in the in total amount of time for which stimulation was given as the pulses were divided into windows. EMG activity was recorded from muscles of the wrist. iTBS resulted in rapid facilitation of synaptic transmission; cTBS resulted initially in facilitation while eventually resulting in inhibition after saturation was reached; imTBS resulted in no net effect. Stimulation was only performed for 40 seconds to 3 minutes, but effects were shown to last well over 60 minutes after the end of stimulation.

Mental health

TMS has also been shown to have many therapeutic for patients with mental health conditions such as depression. A pioneering study was performed by (George et al., 1995) showing that daily rTMS reduces symptoms of depression. Six patients with medication-resistant depression were treated with rTMS in the prefrontal cortex, an area known to exhibit hypofunction. Subjects who experienced depression as a co-morbidity of other mood disorders were treated for at least one week each morning. rTMS was applied at 20 Hz for 20 minutes using a figure-eight coil. For one patient who had been hospitalized several times due to recurrent unipolar depression, rTMS always relieved symptoms of depression and symptoms only relapsed when placebo was given or during unrelated medical illness for which hospitalization was required. Furthermore, glucose metabolism in areas remote from the stimulation site was shown to have returned to normal levels after the 87-day treatment cycle had ended. She reported feeling better than she had felt in three years. In another patient, symptoms such as psychomotor retardation and impaired cognition were present initially but had been eliminated by the end. The study showed that rTMS had promise over traditional drug- and electroconvulsive therapy based treatments for depression. These results were corroborated by (Pascual-Leone et al., 1996), who
performed a more comprehensive study more patients and stricter controls.

Magnetic stimulation of the prefrontal cortex was shown to increase the concentration of dopamine in the striatum of the human brain (Strafella et al., 2001). Eight healthy subjects were administered rTMS at 10 Hz for 30 min. (divided into 3 ten minute blocks consisting of 15 trains of 10 one-second pulses) over the left mid-dorsolateral prefrontal cortex (MDL-PFC) and occipital cortex. The target locations were identified using an MRI, and positron emission tomography (PET) scans with tracers to detect dopamine binding were obtained before and after the experiments. The study showed that rTMS decreased dopamine binding to the left caudate more when the MDL-PFC was stimulated compared to the occipital cortex. Moreover, rTMS did not affect dopamine concentration in the putamen or nucleus accumbens. The exact mechanism by which the dopamine was released was not identified but evidence for both direct modulation by the MDL-PFC and indirect modulation via corticonigral projections were suggested as possibilities and is currently an active area of research (Siebner et al., 2009). Research in this area has a potential for impact in diagnosing and treating Parkinson's, schizophrenia, and addiction.

2.5.2 Interaction with ambient magnetic fields

Over the past few decades, researchers have also described the effects of magnetic fields present in the environment. While most of these studies are in vitro, they could have implications for health. For instance, (Blackman et al., 1985) reported on the release of calcium ions in chicken brain tissue in vitro when exposed to weak extremely low frequency electromagnetic radiation. Specifically, in earth's magnetic field, 15 Hz and 45 Hz electromagnetic waves with a magnetic field strength of 40 Vpp/m were shown to have an effect. The effect was hypothesized to be due to resonance-like effect between the frequencies of stimulation and the strength of the local geomagnetic field (LGF). The authors suggested that experiments performed
on brain tissue in an unshielded environment may be subject to errors because of this effect.

(Azanza, 1989) reviewed a number of findings that suggested that steady magnetic fields of extremely low frequency affect the binding of calcium in the brain. These include stimulating the release of noradrenaline to altering the efficacy of the analgesic morphine. Azana cited the involvement of Ca$^{2+}$ ions in these phenomena and performed a study showing that weak, static magnetic fields of strength 0.116 T and 0.260 T interact with mollusk neurons to inhibit the spiking of bursting cells through a mechanism also involving Ca$^{2+}$ ions. The same neurons silenced by the steady magnetic fields were also shown to be silenced by caffeine. Ca$^{2+}$ was involved in the gating of Na$^+$ ions, and when these ions are removed, the suppression effect was reduced. Thus, the authors suggested that the inhibition was due to changes in ion concentrations which affect the conductivity of ion channels in the neuronal membrane.

(Pacini et al., 1999) showed that neurons in the presence of weak magnetic fields with a strength on the order of 0.2 T experience dramatic morphological changes after only 15 minutes of exposure. Cultures of cells were shown to become abnormally elongated; form vortexes; and/or develop branched neurites with synaptic buttons (typical of cells undergoing plasticity or differentiation). In contrast, non-neuronal cells that were tested - including mouse leukemia and human breast cancer cells - did not show any changes. Furthermore, no changes to the cells' DNA itself were observed. The results suggest that weak static magnetic fields warrant further investigation before being dismissed as safe for long term exposure.

Organisms from homing pigeons to honey bees are thought to utilize magnetic fields for navigation. (Vargas et al., 2006) demonstrated in vitro that neurons extracted from the hippocampus of homing pigeons altered their firing patterns in the presence of changing ambient magnetic fields. Similar findings have been observed
in honey bees (Liang et al., 2016) and sharks (Meyer et al., 2005). A further discussion of an alternative mechanism utilizing magnetite for magnetoception is provided below.

2.6 Biomagnetite and its implications on memory and disease

Biomagnetite is magnetite (Fe$_3$O$_4$) that occurs in living organisms. While some organisms are thought to use magnetite for navigation, for others - including humans - the utility (and source) of biomagnetite is uncertain. We describe here findings related to biomagnetite and how it can impact neuronal activity.

2.6.1 Presence in non-human organisms

(Kirschvink and Gould, 1981) reviewed the presence of magnetite in organisms such as bacteria, sharks, honey bees, and homing pigeons and suggested that magnetite may be responsible for the ability of these organisms to sense very miniscule changes in magnetic field. Various processes ranging from reproductive behavior to navigation are affected by changes in magnetic field. The authors proposed various mechanisms for the biosynthesis of magnetite, not necessarily relating them to the above organisms; for instance, they described that of chitons as a process in which ferric iron stored in the protein ferritin is mineralized through an unknown mechanism and converted into magnetite. According to works cited, ferritin is also present in bacteria and may be mineralized by pigeons. The cubic crystal structure containing tetrahedral and octahedral coordination sites account for its ferrimagnetic property, and a saturating magnetization of about 48 mT is observed for single domain magnetite particles.
Kirschvink et al. reported the first detection of magnetite in the human brain (Kirschvink et al., 1992). The particles are the same ones found in bacteria and fish reported previously (Kirschvink et al., 1992). Single domain crystals tended to clump in groups of 50 to 100 particles, and the densities of the particles within most brain tissues were at least 4 million crystals per gram. Notably, the pia and dura contained densities of over 100 million crystals per gram. Of the patients whose brains were examined, four out of seven had suffered from Alzheimer’s disease (AD). For all brains, tissues were collected from cortical areas and the cerebellum for all brains. SQUID magnetometry was used to measure the magnetic field. It was found that the remnant magnetization of the tissue was approximately 300 mT, which is typical for magnetite but does differ from the results reported by Kirschvink in 1981 (Kirschvink and Gould, 1981). The crystal alignment of the particles found in the study differs from that occurring in geological magnetite, suggesting that the crystals form from a biological process.

The authors stated that two possible biomedical issues may arise as a result of the presence of magnetite in human brain tissue. First, MRI machines can cause the magnetic particles to store 10 to 10,000 times as much energy as that present in covalent bonds, which could pose a health hazard. Second, mechanical oscillations in the particles due to 50 and 60 Hz power line oscillations whose field strength is on the order of earth’s magnetic field (50 µT) could cause ion channels to open; and although most power lines generate fields that are much weaker than this, averaging could result in some effect even at lower levels.

Studies had shown that magnetic fields of only 1-2 mT were enough to induce epileptic seizures in epileptic patients. Based on these findings, (Dunn et al., 1995) investigated the role of hippocampal biogenic magnetite in seizures. Samples were
taken from three deceased normal and epileptic patients and from the removed epileptogenic zone of a living patient and were analyzed for the presence of magnetite particles. The study showed that the particles were more similar in size and structural composition to those found in the GS-15 bacteria, containing a mixture of grains that are magnetically interacting, rather than to MV-1 bacteria, which utilize metabolized single-domain magnetite for magnetotactic behavior (Kirschvink et al., 1992). The opaque particles that were discovered closely match the description of so-called brain sand reported in previous histological studies.

Hautot et al. performed a seminal study in which patients with AD as well as age and sex-matched controls were subjected to SQUID magnetometry (Hautot et al., 2003). They found that patients with AD tended to have significantly more iron in the form of magnetite in their brain tissue. Previous studies had shown that magnetite is capable of producing free radicals, which are known to cause harmful chain reactions in biological tissue (Kirschvink et al., 1992). The brains of three recently deceased subjects who had suffered from AD were placed under deep-freeze in 193 K, ground into a powder, and placed into a small chamber optimized for viewing with a SQUID. Three complementary procedures were performed to separate the magnetite signal from that of other components such the diamagnetic tissue and ferritin (a nonpathological transport protein containing iron). These included measuring the hysteresis loop of the sample, the isothermal remnant magnetization, and magnetization as a function of temperature. Combining the information from these analyses, the authors concluded that patients with AD had between 1 and 7 µg of magnetite per gram of dry tissue while subjects without AD had little to magnetite in their samples. Moreover, the patient with the most severe AD had the highest concentration of magnetite, and the one patient without AD who did have a small amount of magnetite in his sample had shown some signs of neurodegeneration in his autopsy. While the source of magnetite is not known, the preliminary data
suggests it is linked to neurodegenerative disease; detecting magnetite by MRI or MEG could provide an avenue for early diagnosis.

Very recently, Maher et al. confirmed that pathological magnetite was sourced externally, not produced by the body’s metabolism (Maher et al., 2016). A total of 37 brain samples from subjects, some of whom died as a result of AD, in Mexico City and Manchester, UK were analyzed and found to contain high concentrations of magnetite with the same rounded crystal structure as magnetite found in air pollution in urban areas (in contrast to the angular variety found in ferritin). The study involved performing both magnetometry to determine the saturation magnetic remanence, high resolution transmission electron micrography, electron energy loss spectroscopy, and energy dispersive X-ray to determine the composition of the particles. Examination was performed both on tissue slices as well as after tissue digestion by proteolytic enzyme. All samples were found to contain between 0.2 and 12 µg of magnetite per gram of dry tissue. The magnetite particles were found to be intermixed with transition metal nanoparticles, suggesting that they were produced from a process that required intense heating and cooling. The size of the nanoparticles (between 18 and 150 nm in diameter) exceeded that of those found in ferritin (about 8 nm in diameter) by between 2 and 20-fold. Sources could include combustion and/or particles that oxidize when made airborne. In general, the older the age of the subject, the greater the concentration of magnetite, especially in subjects from Mexico City. Magnetite has direct access to the brain via axons in the olfactory nerve, and both AD and healthy subjects’ samples were found to contain magnetite.

2.6.3 Potential for memory storage

(Banaclocha et al., 2010) provided another theory for magnetic memory in the brain suggesting that magnetite may be important for storing information in the cerebral cortex. In this way, magnetite would be essential for perception. Somewhat mir-
roring their cytoarchitectonic theory described above, they hypothesized that the
distribution of magnetite in the brain has a specific form and that the orientation of
single-domain magnetite crystals found in the brain generates magnetic fields that
interact with neuronal activity. The distribution itself may be modified by the fields.
They also suggested that each neuron may have a magnetite on its surface, whose
distribution depends on firing patterns; moreover, the distribution would have an
important role in developing a so-called three-dimensional magnetic construct in
conjunction with astrocytes. This construct would then influence magnetite dis-
tribution on the astroglia themselves, thereby allowing astroglia to store long-term
information. Thus, there would be bi-directional cross-talk between the neurons
and glial cells. Their theory predicts that spontaneous processes such as “creativity,
imagination, thinking, and dreams” could arise as a result of this cross-talk.

Three critiques are presented about this theory. First, from studies on brain mag-
etite, it is known that magnetite forms in clumps that occur in specific areas such
as the hippocampus and superior temporal gyrus (Dunn et al., 1995; Hautot et al.,
2003). By claiming that neurons could have unique magnetite-based signatures, a
notion essential to their theory of spontaneous thought arising from the interaction of
magnetic fields of the magnetite and neuronal activity, Banaclocha et al. contradict
the results of conclusive histological and imaging-based studies showing that mag-
etite is localized in specific areas (Maher et al., 2016). Second, the brains of healthy
individuals who recently passed tended to have very little if any magnetite (Hautot
et al., 2003; Maher et al., 2016). While these individuals were most likely capable
of being creative and thinking, it is likely that brain magnetite was not essential for
these functions. Last but not least, as described above, magnetite is known to cause
the production of free radicals which are highly toxic for living organisms; thus, the
presence of magnetite throughout the brain, as proposed by the theory, could be
deadly (Kirschvink et al., 1992).
Brain-inspired and hybrid analog-digital computing

In the last chapter, we saw how magnetic fields provide essential information about neural signaling which can be detected by methods such as MEG. In the brain, electromagnetic activity allows us to process sensory stimuli and respond with motor commands. This is accomplished by the massive network of neurons which, in a sense, allow the brain to function as a highly efficient computer, performing petaflops of computation while consuming only 12 W of power (Sarpeshkar, 1998).

The incredibly high efficiency of the brain as well as the complexity of the problems it is capable of solving - such as face and speech recognition - have generated interest in research that aims to exploit the properties of the brain to perform pattern recognition in artificial (as opposed to biological) architectures. In this chapter, we briefly overview the field of brain-inspired computing, highlighting two of the most important areas: neural networks and neuromorphic computing. The latter can be implemented in an analog, digital, or hybrid fashion.

This leads us to a recently proposed theory by Nicolelis and Cicurel which suggests that electromagnetic fields in the brain interact with neural activity through a hybrid analog-digital mechanism (Cicurel and Nicolelis, 2015). According to their
theory, magnetic fields are essential for higher level processing that takes place in the brain, such as the real-time integration of sensory information. Along with our findings about hybrid computers, their theory provided the inspiration for our device discussed in subsequent chapters.

At the end, we outline our hypothesis that a hybrid computing device involving magnetic fields is capable of performing nonlinear pattern recognition by utilizing unique features generated as a result of the interaction of magnetic fields in a non-linear medium.

3.1 Neural networks and neuromorphic computing

Artificial neural networks (ANNs) - or simply, neural networks for short - are pattern recognition algorithms loosely inspired by their biological counterparts (Goodfellow et al., 2016). The output of one neuron, or unit, serves as an input to another, and units are arranged in layers which compute in parallel. ANNs are typically not designed to model the brain; rather they operate under the assumption that reverse engineering the brain by breaking it down into its constituents would allow one to replicate the functions it can perform. The constituents considered are only a subset, however; the behavior of ion channels, electromagnetic fields, and many other properties of real neural tissue are not typically represented in ANNs. Despite this, ANNs have been highly successful in accomplishing many classical artificial intelligence (AI) tasks, making ANNs an important brain-inspired technology worth considering here. We discuss the theoretical as well as practical sides, the latter of which is comprised of software and hardware implementations.

3.1.1 Historical background

The first artificial neuron / unit called a linear threshold unit (LTU) was developed in 1943 by McCulloch and Pitts and was so named because a linear function of the
inputs to the neuron is subjected to a threshold to determine its output (McCulloch and Pitts, 1943). The LTU was modeled after the three major parts of a neuron: the dendrites (which receive the inputs), soma (which perform the weighted summation or integration), and axon (which carries the output). No learning algorithm was proposed for the weights of the LTU, which were set manually; however, in 1957, over a decade later, Rosenblatt invented the perceptron, an adaptive version of the LTU that could be trained using the perceptron learning algorithm (PLA). More details about the mechanics of the PLA are given in chapter 5. A key feature of the perceptron is that arranging these units into multiple layers (a so-called multilayer perceptron or feedforward neural network) would allow for classification of nonlinearly separable data. Werbos showed in 1974 that such a network could be trained using an algorithm known as error backpropagation (Werbos, 1974). However, Werbos’ work was mostly overlooked, and backpropagation did not become popularized until the 1980s.

After experiencing an initial surge of interest in the 1960s and early ’70s, ANN research fell into a period known as the first AI winter (Russell and Norvig, 2002; Goodfellow et al., 2016). This period lasted from 1974-1980 and was instigated by misinterpretations of the work of Minsky and Papert who stated that (1) single-layer perceptrons were not capable of learning nonlinear mappings and (2) training large networks of single-layer perceptrons would be too computationally intensive for practical purposes. Nearly two decades of progress that showed promise in areas such as reasoning, natural language processing, and computer vision came to a halt as funding was sharply reduced with investors becoming highly skeptical when reality did not meet expectations.

network consists of bidirectionally fully-connected LTUs, whose outputs are +1 when the weighted input is greater than a threshold and −1 otherwise. In other words, in a system of \(N\) LTUs, there are \(\frac{N(N-1)}{2}\) weights, where the weights are constrained such that (1) the weight \(w_{ij}\) connecting unit \(i\) to unit \(j\) is equal to the weight \(w_{ji}\) connecting unit \(j\) to unit \(i\) and (2) no unit is connected to itself. Hopfield was able to show that such a network always converges to attractor states when driven by a set of \(N\) inputs (i.e. the system is not periodic or chaotic). In the learning phase, the weights are adjusted according to the well-known Hebbian learning rule: units that fire together wire together, and units that fire out of sync don’t link. (Assuming a mean output of zero for each neuron, this amounts to performing a correlation between all pairs of neurons, whose outputs are binary.) In the prediction phase, the weights are fixed, and the output of the network converges to the closest attractor state for that input. Hence, Hopfield networks are based on the brain’s property of associative memory. Rumelhart and Hinton at the same time also rediscovered the backpropagation algorithm, showing that it could be used to efficiently train multilayer ANNs (Rumelhart et al., 1986). Despite these successes, a second AI winter took place in the late 1980s and early 1990s primarily due to the rise of desktop computing.

Advances in computer hardware then provided scientists with the tools necessary to test backpropagation in larger networks. Since then, neural networks have demonstrated state-of-the-art performance in problems ranging from handwritten digit recognition and facial recognition (computer vision) to text translation and speech recognition (natural language processing). The units that comprise these networks employ continuous, differentiable nonlinear transfer functions (rather than thresholds), are trained via backpropagation, and are linked in various architectures which regularize the network by determining which parameters are shared and how they interact with the input.
3.1.2 Architectures

In this section, we briefly review some of today’s most important neural network architectures and describe how they relate to processes in the brain.

Feedforward neural networks and backpropagation

Feedforward networks, or multilayer perceptrons, are inspired by the parallel distributed nature of biological synapses; i.e. a single neuron may be synaptically connected to multiple other neurons which in turn are connected to another set of neurons, and so on (Goodfellow et al., 2016). They have been applied widely to problems in pattern recognition. Figure 3.1 shows a diagram of a standard feedforward neural network. The first layer is known as the input layer. This consists of $D$ nodes whose values comprise a set of $D$ inputs. The input layer directly feeds into a hidden layer. Standard feedforward neural networks are “fully connected,” so all units in one layer are connected to all units in the next. Units in a hidden layer perform a nonlinear transformation of a predefined functional form (e.g., rectified linear or logistic sigmoid) on a weighted sum of their inputs. For example, the output $y$ of a

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{standard_feedforward_network.png}
\caption{Diagram of a standard feedforward neural network with an input layer, two hidden layers, and an output layer.}
\end{figure}
unit with a rectified linear nonlinearity with weights $w$ is given by:

$$y(x; w) = \begin{cases} 
0, & w^T x \leq 0 \\
w^T x, & w^T x > 0
\end{cases}$$

where the last element of $x$ is fixed to 1 so that the last element of $w$ is a threshold. The rectifying function has a support of $[0, \infty)$. Outputs of one hidden layer are treated as inputs to the next. Finally, the output layer performs a nonlinear transformation to generate the prediction for the data point. For instance, when the goal is regression, the output layer nonlinearity may be another rectified linear transformation; when the goal is classification, it may be logistic sigmoid. The output $y$ of a logistic sigmoid unit with weights $w$ is given by:

$$y(x; w) = \frac{1}{1 + \exp(-w^T x)}$$

The function has a support $(0, 1)$. To evaluate the quality of the prediction from any neural network, a loss function must be chosen. Typically, for regression, this is the squared error loss while for classification, it is the cross-entropy loss. Mathematically, the former is given by:

$$L_1(\hat{y}; y) = (\hat{y} - y)^2$$

while the latter is given by:

$$L_2(\hat{y}; y) = -y \log(\hat{y}) - (1 - y) \log(1 - \hat{y})$$

where, in both cases, $\hat{y}$ is the prediction. This loss is usually averaged over all data points. Assuming $y \in (-\infty, \infty)$, we find that $L_1, L_2 \in [0, \infty)$.

To adjust the weights such that the loss function is minimized, we solve for when the derivative of the loss is zero. When the network is composed of many layers and various nonlinearities, writing this expression and solving for the minimum becomes challenging. Backpropagation simplifies this process by allowing one to
write the derivative of the loss in one layer with respect to (1) the weights in that layer and (2) the outputs of the previous layer. This process is repeated recursively backwards from the output layer to the input layer, resulting in simple expressions that allow the weights to be updated methodically one layer at a time. For a long time, backpropagation was thought not to have any biological analogs. However, recently, researchers have proposed a number of ways that backpropagation could be implemented by the brain; for a detailed description, see (Marblestone et al., 2016).

Convolutional neural networks

Convolutional neural networks (CNNs) are inspired by the human visual cortex which is known to contain layers that each process more and more abstract features of a visual scene. For instance, neurons in the lowest layer might be responsive to edges, the layer above to shapes, the layer above to faces, and so on. Neurons in these layers contain so-called receptive fields, which are filters that search for features of interest such as those listed above. The filters are invariant to the locations of their targets within a scene, and each neuron may be sensitive to more than one interacting receptive field.
CNNs are based off of this phenomenon and are commonly employed to problems in image processing such as object / facial recognition and self-driving cars. Figure 3.2 shows a diagram of a CNN architecture (Raghav, 2018). Unlike the feedforward network, the units in a layer of a CNN are arranged rectangularly to correspond with the pixels of an image. The depth of each layer signifies the number of inputs per pixel (input layer) or the number of receptive fields acting on the units in the previous layer (hidden layers). For example, if a hidden layer has a depth of five, then there were five receptive fields applied to the previous layer. Because the pixels can be color, the input layer has a depth of 3, corresponding to red, green, and blue intensities. For any layer, receptive fields, which contain the weights for the CNN, are convolved across units in the layer to produce an output. These outputs are then fed element-wise through a nonlinearity such as the rectified linear or logistic sigmoid described above. The size of the receptive field and features like pooling affect the size of the subsequent layer. (Note that the receptive field is usually much smaller than the size of the image and the same receptive field is convolved across the entire image: due to this weight sharing, the number of weights in a CNN is much less than that of a feedforward network with the same number of inputs.) After a number of convolutional layers, a few feedforward layers are usually appended to perform the classification itself.

Training a CNN is often also done with backpropagation. This involves first backpropagating errors through the feedforward layers, and then, backpropagating through the convolutional layers. The latter requires recognizing that the same weights in a receptive field that act on one part of an image also act on another part. Mathematically, the procedure involves generating a gradient matrix and performing another convolution - this time with a rotated matrix - to get the updated weights.
Recurrent neural networks

Recurrent neural networks (RNNs) are a third major class of neural network architectures often used for processing time-series data (Goodfellow et al., 2016). They were inspired by feedback loops present throughout the brain and nervous system. Figure 3.3 shows a block diagram for a generic RNN (Britz, 2015). For each time point, a data point $x_t$ is inputted. The hidden state $h_t$ is a function of $x_t$ as well as the hidden state at the previous time point $h_{t-1}$, hence the term “recurrent” in the title. Finally, the output $o_t$ is a function of only the hidden state $h_t$. The standard RNN restricts the parameters that model the hidden state and the output to be fixed; hence, the number of parameters to be learned in an RNN is typically much fewer than that of a feedforward network. Learning is done using a variant of backpropagation known as backpropagation through time (Werbos, 1990). Both in terms of prediction and learning, the basic RNN shares many traits with the Kalman filter (and its extensions), a model commonly used for brain-machine interface decoding.

One of the most popular variants of the standard RNN is the long short-term memory (LSTM) RNN (Goodfellow et al., 2016). The LSTM retains information
that was presented earlier on in the time series in addition to in the recent past. It contains three types of “gates” - the “forget” gate, the “input” gate, and the “output” gate - as well as a parameter that represents the current state. The latter is a higher level feature of the LSTM that allows it to retain long-term memory. The forget gate decides which information from the current state should be thrown away based on the input and previous hidden state; the input gate combines the previous state and current input to determine what the new state should be; and the output gate generates an output based on previous hidden state and current input. LSTMs have been employed widely on time series data ranging from stock market prediction to sentence completion to video interpolation (Goodfellow et al., 2016). In the case of sentence completion - in which the first few words are inputted as a query to generate the last - context may be important; without an LSTM, information that was presented much earlier in a sentence or paragraph (such as the name of a place or a person’s gender) might be lost.

3.1.3 Implementations

In the last section, we reviewed the basic architectures for performing generic pattern recognition, computer vision, and time series analysis. The mathematical equations that govern neural networks must be implemented in some type of medium - be it software, hardware, of physical (e.g. mechanical, acoustic, optical, etc.) - in order to be useful beyond theoretical analysis. In this section, we review various implementations of neural networks and assess their tradeoffs. Since the device described in this dissertation is hardware that utilizes the interaction of magnetic fields in a physical medium, we will focus our attention on the latter two cases.
Software

Neural networks are most commonly implemented in software that runs on digital hardware. Over the past few years, many open source software packages including TensorFlow, Theano, Caffe, and Torch have been developed enabling individuals with relatively modest personal computers to create and deploy neural network architectures for their machine learning applications (Garbade, 2018). Together with today’s powerful and affordable graphics processing units (GPUs) - which are optimized for the highly parallelizable matrix operations used to train and predict using a neural network - these software packages have made software implementations not only convenient but extremely powerful (NVIDIA, 2018). Lastly, software offers great flexibility since making modifications requires only changing lines of code and recompiling; i.e. physical changes such as rewiring are not required.

Hardware: analog, digital, and hybrid

Hardware implementations of neural networks are known as neuromorphic chips; to various degrees, these are designed to mimic the behavior of biological neurons. Most neuromorphic chips are based on very large-scale integration (VLSI) technology. VLSI includes hardware in which thousands to billions of transistors or memristors are embedded onto a single wafer of silicon. This allows for reduced power consumption and latency since long wires are not needed to connect the different components of a computer to each other. VLSI itself is divided into two major categories: analog and digital. The majority of neuromorphic chips are either analog VLSI, hybrid analog-digital VLSI, or hybrid designs involving memristors.

Analog circuits involve voltages which are continuous in both amplitude and time. They are typically designed using transistors operated in the subthreshold region. (Arima et al., 1991) showed that they could design a neural network consisting of 336 neurons and 28224 synapses using complementary metal-oxide semiconductor
(CMOS) transistors. Over $10^{12}$ computations per second are performed per chip with the capability to scale to hundreds of chips. The network is “self-learning,” meaning that network reorganizes itself after being presented with inputs, similar to a Hopfield network. (Alspector et al., 1992) show that it is possible to create a Boltzmann machine - a stochastic variant of a Hopfield network - in analog VLSI. They show they are able to perform classification for the generalized XOR problem with up to eight input bits using a network of 32 neurons with 496 synapses at a speed of $10^8$ computations per second. (Shima et al., 1992) show one of the first VLSI designs which are capable of implementing both Hebbian learning (similar to the above two schemes) as well as backpropagation (today’s state-of-the-art). (Bo et al., 1999) demonstrated a major improvement to the analog backpropagation algorithm by allowing the learning rate to be adaptive, leading to faster convergence. (Farquhar et al., 2006) introduced the field programmable neural array, which consisted of analog building blocks that could be easily reconfigured and combined into complex designs. The goal was to model neurons in hardware down to the individual ion channel.

On the other hand, digital circuits involve discrete values and are typically composed of transistors that are operated in the saturation or cutoff regions. Compared to analog, they are easier to design, offer a higher signal-to-noise ratio, and can be easily combined into larger circuits (Beiu, 1997). (White and Elmasry, 1992) implemented the neocognitron, the precursor to CNNs which were based off the discoveres of Hubel and Wiesel, in digital VLSI. (Botros and Abdul-Aziz, 1994) demonstrated one of the first implementations of a feedforward neural network using field programmable gate arrays. Their network consisted of only three layers with 5 input neurons, 4 hidden neurons, and 2 output neurons, and training was also performed offline on a computer.

Hybrid neuromorphic architectures involving analog and digital VLSI have also
been proposed and implemented. Hybrid computing involves combining analog and
digital computing elements to get the best of both. In general, analog computers
are faster but less accurate while digital ones are limited by clock speed and require
inputs and outputs to be quantized; the latter is what makes them more accurate.
(Murray et al., 1991) proposed a pulse-stream neural network in which neural states
are represented as spike trains. This consisted of 15000 synapses which allow a robot
to navigate in a $24 \times 24$ grid environment. (Fang et al., 1992) were able to implement
a hybrid architecture for image compression: the analog component allowed the
designers to minimize surface area and power consumption while the digital part
was used to store the weights. (Coggins et al., 1995) used a similar architecture with
three layers and 180 weights for intracardiac morphology classification.

Memristors are devices whose resistance depends on the amount of charge that
has gone through them. They were first discovered by scientists at Hewlett Packard in
2008 (Strukov et al., 2008). Because of their natural ability to store information, they
have been intensely investigated by the neuromorphic computing community in the
past decade. Many memristor-based devices follow a hybrid architecture, in which
analog and digital CMOS transistors are combined with memristor. (Affi et al.,
2009) proposed a hybrid CMOS-memristor crossbar which could implement Hebbian
learning. Information could be sent bidirectionally along axons and dendrites, al-
lowing for local learning. (Jo et al., 2010) built one of the first hybrid nanodevices
consisting of both memristors and standard CMOS transistors. They fabricated a
device which has a synaptic density of $10^{10}$ connections per cm$^2$ and show that it
is also capable of performing Hebbian learning. (Indiveri et al., 2013) were the first
to propose a hybrid memristor-CMOS circuit that modeled individual ion channels,
not just synapses. This would allow the circuit to behave probabilistically and make
it more fault tolerant. (Volos et al., 2015) showed that it is possible to couple two
neuromorphic circuits using a magnetic flux-based memristive device. Evolution of
the coupled system shows complex dynamics including periodicity, synchronization, and chaos. (Prezioso et al., 2015) improved on the work of (Afifi et al., 2009) by avoiding the need for a transistor at each location on a crossbar; this greatly improves scalability. They demonstrated that they were able to construct a perceptron to perform linear classification of $3 \times 3$ images into three categories. Most recently, (Wang et al., 2017) developed a memristive device that exhibited temporal dynamics of Ca$^{2+}$ channels and verified by microscopy that Ag particles used in the device disperse and regroup in a pattern similar to that of biological ion channels.

Physical

Aside from the analog and digital electronic architectures that have been developed to train and predict with neural networks, forays have been made into building neural networks into physical media. As described by (Hermans et al., 2015), forward propagation in reservoir computing networks - those for which a very large number of random hidden features are generated by the inputs but which are left untrained with the assumption that at least one of the features will allow data to be made linearly separable - has been implemented in "water ripples, mechanical constructs, electro-optical devices, fully optical devices, and nanophotonic circuits." Thus, we restrict our discussion here to circuits which implement backpropagation in analog physical media.

(Hermans et al., 2015) showed that an acoustic medium that consists of a source and receiver can be turned into any type of neural network that implements backpropagation. Their system consists of a set of sources and receivers at identical locations; half the time, half the sources are active and vice versa. When the source at a particular location is active, the receiver is not. They choose a rectified linear unit (ReLU) nonlinearity for their demonstration, which has a simple derivative with
respect to errors and weights:

\[
\frac{\partial \text{ReLU}(x)}{\partial x} = \begin{cases} 
0, & x \leq 0 \\
1, & x > 0
\end{cases}
\]

This simplifies multiplication that is involved in backpropagation. A similar setup is employed for their electro-optical setup in which a grid of lights drives a set of neurons which then feed back onto the grid (all through fiber optic cables).

While the networks created by Hermans et al. consisted only of a handful of neurons, (Shen et al., 2017) showed that an optical network consisting of 56 Mach-Zehnder interferometers can be used for a vowel recognition task. Because the architecture is entirely optical, matrix multiplication is performed optically, resulting in dramatic speed and power efficiencies over electro-optical methods. Among the main limitations of the network are thermal crosstalk between interferometers and the precision of optical phase. Their device is engineered to employ the finite difference method to perform gradient descent rather than backpropagation. (Hughes et al., 2018) and (Lin et al., 2018) later introduced methods to perform backpropagation in purely photonic circuits and applied them to problems in computer vision.

3.2 Hybrid analog-digital theory of neural computation

In the previous section, we discussed neural networks and their various implementations in analog and digital hardware and physical media. We found that hybrid neuromorphic architectures were often more efficient in terms of space and power consumption than their analog or digital counterparts, and moreover, physical media allowed for learning to take place in a network in real-time. These ideas are closely tied to a theory proposed in (Cicurel and Nicolelis, 2015). Here, we outline the basic tenets of that theory and state the hypothesis which governs the design of the device described in this thesis.
3.2.1 The Relativistic Brain

The central idea behind the relativistic brain theory (RBT) is that neurons in the brain generate electromagnetic fields which form a manifold the authors call the neuronal space-time continuum. This field stores information and then feeds back onto the neurons themselves. The interaction between the analog fields and the digital spikes that generated them make the brain a hybrid analog-digital system. The authors claim that because the fields are represented in analog, they, and the brain itself, cannot be simulated by a digital computer. Thus, no Turing machine - a description for any computer which can follow a set of instructions and stores information in addressable memory - can simulate the brain.

One of the key observations the authors make that suggest that electromagnetic fields are involved in neural computation is called the “degeneracy principle” (Cicurel and Nicolelis, 2015). According to this principle, the exact same neurons need not fire to elicit a certain motor response. Moreover, the brain being a stochastic and/or chaotic system also may not reliably produce the same exact spiking pattern when it is presented with same stimuli. These findings were validated with experimental work involving brain-machine interfaces. The authors suggest that neuroelectromagnetic fields (NEMFs) act as a “glue,” integrating information from neurons in the vicinity. The fields store information in analog that are produced by the digital spikes, and the same field at a given location can be generated by multiple patterns of neural activity, which explains the degeneracy principle. Another key observation is that the NEMFs are generated in major white matter bundles in the brain, many of which form loops. While fields due to individual neurons may be weak, the combined field strength of many neurons within a bundle may be significant.

The NEMFs also give the brain a “sense of self” (Cicurel and Nicolelis, 2015). This allows the brain to (1) perceive complex visual scenes, (2) anticipate events,
and (3) incorporate external sensations / limbs as the body’s own. Disruption of these fields via TMS can lead to changes in perception and ability to utilize the senses; furthermore, mental illnesses are hypothesized to be a direct result of pathological NEMFs. Thus, the fields allow us to integrate information from the environment rapidly, something a purely digital mechanism would struggle to do. Hence, the authors claim that the brain cannot be simulated on a purely digital computer.

The RBT also suggests that the brain cannot be simulated by a Turing machine (Cicurel and Nicolelis, 2015). The authors state that the Church-Turing hypothesis which claims that, “any function which would be naturally considered as ‘computable’ could be computed by a universal Turing machine,” does not apply to the brain. This is because information represented in the brain is physical rather than syntactic; NEMFs provide the brain with emergent properties that cannot be encoded syntactically as in a Turing machine. Treating any measurable quantity in the brain - all the way from macroscopic electromagnetic fields down to quantum level phenomena - in isolation would bias any digital simulation because it would be unaware of the bulk effect. Moreover, with only a limited computational power and number of sensors to measure brain activity, digital simulations can only yield an approximation. Since the dynamics of neural activity are very sensitive to initial conditions, digital simulations will also quickly diverge from the truth. (A similar problem is faced when attempting to digitally simulate protein folding or forecast the weather; there are not enough measurements currently made to capture the physics and bulk effect that takes place during the processes.) Thus, the brain is not Turing-computable.

3.2.2 Hypothesis

Now that we have now shown the advantages of hybrid analog-digital computation and explained the relativistic brain theory posited by Cicurel and Nicolelis, we come to the hypothesis of this thesis. We claim that analog magnetic fields can be made
to interact in a medium that results in nonlinear features being generated from the inputs. Furthermore, these features can be employed in a digital system to perform classification using a linear classifier on an otherwise non-linearly separable dataset. We aim to show by proof of concept that unique features generated by analog magnetic field interactions are necessary for complex pattern recognition problems, allowing us to draw connections between our device and the brain.
Analog computation: the neuromagnetic reactor

Inspired by our knowledge of the role of synaptic activity and biomagnetic fields in neural signal processing, we have developed a new brain-inspired hybrid computational architecture. In this chapter, we describe the method by which the analog component of this hybrid architecture was constructed and experiments performed to validate its functionality. Our analog computer was designed to physically resemble the network of white matter pathways in the brain. Thus, we begin by describing white matter tractography, a technique used to map out white matter pathways which we employed to design the physical structure of the device. As part of our theory, magnetic fields generated by white matter are hypothesized to influence computation performed by the brain; hence, in our device, magnetic fields play the role of the analog computer. We then describe how the analog magnetic fields interact nonlinearly to produce features that are inputted into a digital component inspired by synaptic activity of neurons.
4.1 Diffusion weighted imaging

The orientations of several major white matter pathways in the brain define the physical structure of our device. To determine how these pathways are distributed in the brain, we perform tractography on a set of diffusion weighted images (DWIs). Diffusion weighted imaging (also DWI) is a type of MRI in which the direction of white matter growth can be inferred based on the diffusion of water molecules. In gray matter, diffusion is isotropic; however, diffusion in white matter is anisotropic, and water preferentially diffuses along the direction in which an axon grows. Diffusion MRIs are formed by a series of volumetric images that are taken by applying magnetic gradients in many directions in three-dimensional space. For each direction, the intensity of a given voxel tells how much water diffused along that direction.

Based on this information, we can build a $3 \times 3$ tensor of values which shows the intensity of diffusion along three primary axes:

$$
\mathbf{D} = \begin{bmatrix}
d_{11} & d_{12} & d_{13} \\
d_{21} & d_{22} & d_{23} \\
d_{31} & d_{32} & d_{33}
\end{bmatrix}
$$

This matrix is symmetric and positive semidefinite, resembling a correlation / covariance matrix that is often encountered in the field of statistics. In fact, to build this tensor, we could simply compute an empirical correlation matrix using the sampled diffusion directions and their intensities. For example, if we measured diffusion in $N$ directions $\mathbf{v}_1, \ldots, \mathbf{v}_N : \mathbf{v}_i \in \mathbb{R}^3$, and recorded the intensities to be $\theta_1, \ldots, \theta_N : \theta_i \in \mathbb{R}^+$, we could construct an empirical covariance matrix $\mathbf{\hat{D}}$ by performing $\mathbf{V}^\top \mathbf{\Theta} \mathbf{\Theta}^\top \mathbf{V}$, where

$$
\mathbf{V} = \begin{bmatrix}
v_{11} & v_{12} & v_{13} \\
v_{21} & v_{22} & v_{23} \\
\vdots & \vdots & \vdots \\
v_{N1} & v_{N2} & v_{N3}
\end{bmatrix}
$$
is the design matrix in which row $i$ is $v_i^T$ and

$$
\Theta = \begin{bmatrix}
\theta_1 & 0 & \cdots & 0 \\
0 & \theta_2 & \cdots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \cdots & \theta_N
\end{bmatrix}
$$

is the diagonal matrix of intensities. Since the basis vectors are unit vectors, the covariance matrix is equivalent to the correlation matrix.

Diagonalizing this tensor yields an estimate of the primary diffusion directions as well as the intensity of diffusion in those directions

$$
\hat{D} = U\Lambda U^T
$$

where the columns of $U$ are the primary diffusion directions, or eigenvectors, and the values along the diagonal matrix $\Lambda$ contain the corresponding intensities, or eigenvalues. The closer the eigenvalues are to each other, the more isotropic the diffusion; on the other hand, if one eigenvalue predominates, diffusion is anisotropic.

(This is a very simple model of the diffusion tensor explained for illustrative purposes. In reality, a single voxel could contain fibers that extend in multiple directions, and hence, more complex models that capture mixtures of diffusion directions, such as the ball-and-stick model discussed below, are often employed.)

DWIs were collected from a healthy male subject (age 30, weight 76.5 kg) by Dr. Nouchine Hadjikhani’s laboratory at Massachusetts General Hospital, Cambridge, MA. The images were acquired at $x$-, $y$-, and $z$-sampling resolutions of 2 mm. In addition, 1 mm resolution T1-weighted images were also acquired to distinguish anatomical features in the brain. Figure 4.1 shows diffusion tensor images (DTIs), which is a visual representation of diffusion tensors that is derived from DWIs. Red corresponds to medial-lateral fiber growth, blue to superior-inferior, and green to anterior-posterior. From the sagittal view (figure 4.1a), the corpus callosum is apparent. Similarly, in the coronal view (figure 4.1b), we see the corticospinal tract,
Figure 4.1: Diffusion tensor images. Colors indicate the primary direction of axon growth in each voxel. Red corresponds to growth in the medial-lateral direction, blue to superior-inferior, and green to anterior-posterior. In figure 4.1a, we see the corpus callosum highlighted in red. In figure 4.1b, the corticospinal tract is highlighted in blue and the corpus callosum again appears in red. Finally, in figure 4.1c, the superior longitudinal fasciculus appears in green and the corticospinal tract in red.

and in the axial view (figure 4.1c), we see the superior longitudinal fasciculus and the forceps major and minor of the corpus callosum.

4.2 Tractography

Tractographies are maps of white matter that are inferred from DTIs (Basser et al., 2000). Tractographies typically have two main purposes: either to perform an exploratory analysis for discovering to which other regions of the brain a particular region is connected or for discovering the shape of a known pathway within the brain (Yendiki, 2010). The former is called local tractography since it considers voxels along a path sequentially (i.e. there are no constraints on where the tracts should end) while the latter is called global tractography since it considers all imaged voxels simultaneously (i.e. tracts should begin and end at specified locations and only the
path between the points must be inferred). Tractography can also either be deterministic or probabilistic; i.e. only the most likely direction of diffusion is chosen at each voxel or the entire probability distribution of possible directions is used to generate a number of possible pathways.

Tractography requires two major preprocessing steps: registration (alignment) and fitting diffusion tensors (Yendiki, 2010). First, T1-weighted anatomical MRIs are registered to a template image to ensure that anatomical structures are placed in correspondence (to account for differences in brain size and orientation). Next, DWIs are registered to the anatomical images to ensure that white matter tracts that are identified are oriented in the right direction with respect to the anatomy of a subject. After registration, one of a number of different models - such as the diffusion spectrum, orientation distribution function (i.e. ODF or Q-ball), ball-and-stick, or basic tensor - must be fit to the DTI to infer the tracts.

In this study, we utilize the software TRACULA (TRAacts Constrained by UnderLying Anatomy, part of the Freesurfer MRI data analysis package) to fit a tractography (Yendiki et al., 2011). TRACULA first performs a cortical reconstruction to identify salient structures in the brain. These structures are used to regularize, or constrain, the inferred tractography. Reconstruction involves linear registration, in which the same deformation is applied to every voxel in an image such that edges in an image best correspond with edges in a template image such as the MNI 152 (Jenkinson and Smith, 2001; Jenkinson et al., 2002). The MNI 152 is a standard template image produced by averaging T1 data from 152 adult subjects; the reference frame is commonly used in the DTI literature (Fonov et al., 2009, 2011). Then, tractography is performed via Bayesian Markov chain Monte Carlo (MCMC) sampling-based inference. The algorithm combines information from the DWI with anatomy inferred from cortical reconstruction of MRIs and returns distributions of pathways between major structures in the brain. During this procedure, TRACULA utilizes a so-called
ball-and-stick model of the diffusion tensor at each voxel. The ball-and-stick model fits a mixture of one isotropic component and up to $N$ anisotropic components, also known as crossing fibers, to each voxel. Crossing fibers can intersect at a voxel and can lead to inaccurate estimates of the diffusion tensor: if a point estimate of the tensor is employed, it would represent just the average diffusion direction even if diffusion in multiple directions occurs with equal probability. TRACULA accounts for this by utilizing a technique called BEDPOSTX (Bayesian Estimation of Diffusion Parameters Obtained using Sampling Techniques, and X indicates crossing fibers) (Behrens et al., 2003). BEDPOSTX fits the ball-and-stick model by performing MCMC sampling to build a full distribution of possible fiber orientations at each voxel. Subsequently, a function called PROBTRACKX is employed to build a probabilistic tractography from the voxel-wise distributions by mapping the most likely pathway over all voxels. Here, we choose $N = 2$.

Compared to other state of the art methods, TRACULA presents two main advantages. First, as its name implies, TRACULA regularizes the inferred tractography by preferring paths that are more anatomically feasible. This is accomplished by a two-fold method. First, ROIs are automatically selected by registering a subject’s anatomical MRI with a template MRI such as the MNI 152. This allows users to avoid having to hand-label individual voxels as belonging to a certain anatomical feature. Second, TRACULA performs global tractography, so only pathways that join ROIs are considered feasible.

Figure 4.2 shows the output of TRACULA. TRACULA identifies a total of 18 tracts. A list of these tracts along with the colors they correspond to in figures 4.2a and 4.2b is given in figure 4.2c. A number of these tracts form large loop-like structures. For example, the forceps major and forceps minor of the corpus callosum extend from one hemisphere to the other in a circular pattern. The superior portion of the corticospinal tract forms part of a loop as a branches out from the inferior
Figure 4.2: Output of TRACULA. 18 major white matter pathways are identified. The shapes of these pathways inform the physical structure of the magnetic reactor. The thickness of the tracts depict a 20% highest posterior density confidence region among all sampled pathways.

region. The parietal and temporal superior longitudinal fasciculi also curl downwards on the lateral sides of both hemispheres. The orientations of these tracts inspired the geometry of the analog part of our hybrid computer described below.

4.3 Large-scale endogenous cortical magnetic and electric fields

Upon observing the loop-like structures in our tractography studies, we began to wonder whether the loops had some functional significance aside from joining the
regions of the brain synaptically. Perhaps these loops could act as transformers, inductively coupling electromagnetic fields generated by cortical and subcortical activity. As a first pass, we decided to investigate the magnitude of the fields that might be produced in comparison to transcranial magnetic stimulation (TMS), which is a noninvasive stimulation technique known to depolarize neurons. TMS has been employed to treat various conditions ranging from psychiatric disorders to loss of nerve function (Janicak and Dokucu, 2015; Khedr et al., 2015). The basic operating principle behind TMS is one of Maxwell’s equations known as Faraday’s law, which states that a time-varying magnetic field produces a curling electric field (Griffiths and College, 1999). If this field is generated in a conductive medium, the voltage difference between neighboring points will create a current, known as an eddy current. TMS induces eddy currents in the brain and spinal cord to diagnose and treat neurological disorders. An external magnetic field is generated by a circular coil carrying hundreds to thousands of amperes of current (Barker et al., 1985). The field is focused on a small brain area, creating a magnetic flux that then generates an electric field capable of exciting neurons.

To determine whether large-scale intracortical magnetic fields could have an effect similar to that of TMS, we performed calculations of the strength of eddy currents that would be induced by current-carrying axons. We based our calculations on similar ones performed to approximate the field strength of a TMS coil. We first calculated the electric field strength of the eddy currents induced by the magnetic field generated by a TMS coil. Typical half-value depths and tangential spreads of TMS range from 0.9 - 3.4 cm and 5 cm$^2$ to 34 cm$^2$, respectively (Deng et al., 2013). Assuming a circular focal region with radius $r$ of 3.5 cm, that the magnetic field strength $B$ of the coil is 2 T, and the flux takes place over 300 $\mu$s (Ro et al., 1998); from Faraday’s law, we can roughly approximate TMS-induced eddy currents to have
electric fields $\mathbf{E}$ on the order of:

$$\int \mathbf{E} \cdot d\mathbf{l} = -\frac{d}{dt} \int \mathbf{B} \cdot d\mathbf{A}$$

$$\mathbf{E}(2\pi r) = -\frac{d}{dt} B(\pi r^2)$$

$$\mathbf{E} = -\frac{r}{2} \frac{d}{dt} \mathbf{B}$$

$$\approx 116 \text{ mV/mm}$$

This agrees roughly with values simulated in more sophisticated models, which fall in the range of 200 to 300 mV/mm (Salinas et al., 2009). Thus, the magnitude of electric fields generated by TMS surpasses what is typically needed to induce action potentials by about 50-fold (approximately 10 V/m (Bonmassar et al., 2012)).

If we model a single neuron as a loop of wire, we can roughly estimate the magnetic field produced by a single neuron. Assume the axial resistance (along the membrane of an axon) is equal to the resistance outside an axon (which is approximately true for a fiber tract). Using cable equations, we find that the forward axial current (due to voltage propagation) is $I_a = -\frac{d}{dt}(V_m)/(r_a + r_e)$, where $V_m$ is the membrane voltage, $\frac{d}{dt}$ is the time-derivative of membrane voltage, and $r_a$ and $r_e$ are the axial and extracellular resistances (Malmivuo and Plonsey, 1995). Peak axial current is on the order of nanoamps for neurons with a resting membrane potential of -70 mV. Define $\mu = \mu_0 = 4\pi \times 10^{-7} \text{ H/m}$ to be the magnetic permeability of neural tissue (Malmivuo and Plonsey, 1995). Assume the loop has a radius of $d = 3.5 \text{ cm}$ (as in the case of the TMS coil above) and that $I_a$ is the current through the loop ($\approx 10^{-9} \text{ A}$).

By the Biot-Savart law (Griffiths and College, 1999), we find that

$$\mathbf{B} = \frac{\mu I_a}{2d} \approx 1.8 \times 10^{-14} \text{ T}$$

at the center of the loop. Assuming the change in magnetic field happens over 1 ms, the approximate duration of an action potential (Malmivuo and Plonsey, 1995), we
can calculate the electric field along points tangent to the neuron as we did above (Griffiths and College, 1999):

\[ E = -\frac{d}{2} \frac{d}{dt} B \approx \pi \times 10^{-13} \text{ mV/mm} \]

Rather than considering the effect only of a single neuron, we are interested in characterizing effects due to large bundles of neurons acting together. The thickest white matter bundle in the human brain, the corpus callosum, contains around 200 million axons (Tomasch, 1954). Treating the corpus callosum as a cylindrical solenoid with this many loops (which all lie on a single plane, a generous approximation), we can calculate the magnetic field strength by the Biot-Savart law (Griffiths and College, 1999):

\[ B = \frac{\mu N I_m}{2d} \approx 3.5 \times 10^{-6} \text{ T} \]

where \( N \) is the number of turns \((200 \times 10^6)\). Then the electric field strength \( E \) at a point infinitely close to the corpus callosum is

\[ E = -\frac{d}{2} \frac{d}{dt} B \approx 2\pi \times 10^{-5} \text{ mV/mm} \]

While this is much weaker than the strength needed for initiating an action potential from baseline (see section 2.4.1), the fields may still be strong enough to modulate membrane potentials, causing neurons that are near threshold to fire. Assuming a 100 \( \mu \)m wide neuron takes about 100 ms to reach a 10 mV spike threshold (Anastassiou et al., 2011), magnetic effects may be able to phase-shift the spike times of neurons by sub-microsecond time delays \((2\pi \times 10^{-5} \text{ mV/mm})(0.1 \text{ mm})(100 \text{ ms})/(10 \text{ mV}) \approx 62 \text{ ns}\). Interactions among fields produced by the billions of neurons in the brain may further augment this effect when they are considered as a whole.
4.4 Neuromagnetic reactor

As described in the chapter 2, endogenous alternating electromagnetic fields can couple neurons, delay spike times, and in some circumstances, cause neurons to depolarize. Furthermore, some have theorized that ferromagnetic nanoparticles in the brain could play a role in signaling. Based on these observations as well as the loop-like structures of white matter observed in our studies of tractography, we created a brain-inspired analog computing device that utilizes magnetic fields for computation. Here, we describe the iterative design of the device, which we call the neuromagnetic reactor.

4.4.1 Design 1: wire model

Our initial goal was to mimic very roughly the shape of the white matter tractography discovered above. To do this, we built a wire model of the tractography about three times the scale of the human brain consisting of wires that modeled major white matter tracts. This initial design is shown in figure 4.3. Our hypothesis was that currents passing through these wires would generate magnetic fields that we could measure using an array of Hall effect sensors (shown in figure 4.3c) and could subsequently use for performing computations such as pattern recognition. Since it is well-known that electromagnetic fields interact linearly in air, we planned to immerse the device in various types of media including saline and ferrofluid to (1) emulate the composition of tissues in the brain and (2) determine whether these media allow the magnetic fields to interact nonlinearly. If the latter were true, medium-induced analog nonlinearities could be utilized to perform pattern recognition similarly to how artificial neural networks perform pattern recognition by learning nonlinear functions that map inputs to outputs (Bishop, 2006). Initially, the nonlinearities could be hand-engineered by tuning the configuration of the tracts, the type of medium, the
Figure 4.3: Design 1: wire model of neuromagnetic reactor. Each of the wires is positioned roughly based on the discovered tractography (shown in figure 4.2). Analog signals are generated from the computer and outputted by a NI-DAQ (PCI-6723) board to a Cerwin-Vega! CV-2800 audio amplifier (National Instruments, 2017; Cerwin-Vega!, 2011). The audio amplifier provides a large current drive and adjustable voltage gain. Currents are fed into different channels of the reactor and magnetic fields are measured by an array of Hall effect sensors (Honeywell, 2015). Each location on the array consists of three sensors, one for measuring the $x$-, $y$-, and $z$-components of the magnetic field, respectively. The sensors are placed on a grid with approximately 19 mm spacing. The magnetic fields are recorded by another NI-DAQ board (USB-6225) and are post-processed on the computer (National Instruments, 2016).
signals inputted, etc., with the eventual goal of allowing the device to adapt on its own.

To begin, we reduced our device down to just a pair of wires and performed experiments in which we passed alternating currents at 50 Hz and 93 Hz to determine the relationship between the applied currents and measured magnetic field. Figure 4.4 depicts the experimental setup and the Hall effect probe. Two 10 gauge wires were fed through a container of diameter 6 cm and were placed about 1 cm height difference apart from each other. These wires are highlighted in blue and green in figure 4.4. The wires were oriented 90 degrees from each other and shielded everywhere except inside the cup to minimize the effects of electromagnetic coupling / induction and to isolate the effect of interaction in a medium. Sinusoidal signals were generated on a computer at a sampling rate of 50 kHz and outputted via a National Instruments (NI) PCI-6723 Data Acquisition (DAQ) analog output module (National Instruments, 2017). These signals were low pass filtered with an RC circuit with a cutoff frequency of 400 Hz and fed to a Cerwin-Vega! CV-2800 audio amplifier capable of generating up to 2800 W of peak power and 700 W continuous power (Cerwin-Vega!, 2011). Three Honeywell SS49E Hall effect sensors were attached orthogonally to one another (assembled into a probe) and placed 1.7 cm laterally away from the crossing point of the wires (depicted as an orange dot in figure 4.4b and on the plane halfway between the two wires (Honeywell, 2015). The sensors were configured to measure the magnetic field in each of the principal directions: $x$, $y$, and $z$. To measure the current through wires, we used a NI USB-6225 analog input module to measure the voltage across a 200 mΩ resistor connected in series with each of the amplifiers and divided by the resistance to get the current (National Instruments, 2016). This current $I$ (units of amperes, or A) is proportional to the magnetic field strength $H$ (units of A/m) since for a straight wire, $H = \frac{I}{2\pi r}$ where $r$ is the distance the Hall effect sensor is away from the wire.
(a) Circuit diagram for single wire experiment

(b) Experimental setup

(c) Hall effect probe

Figure 4.4: Experimental setup to measure magnetic field of a single wire and a pair of wires in various media. In all three media tested, the magnetic field was very weak (sub-mT) and exhibited only linear interaction.

We planned first to verify that the setup was working as expected in air; i.e. that the frequency of the current and that of the magnetic field were the same and that the permeability of air could be recovered by relating the measured magnetic field to the applied magnetic field strength. The latter would indicate that our calculations of magnetic field strength were accurate. In the following experiments, we applied currents $I_1$ and $I_2$ of amplitude approximately 33.6 A and 18.1 A for four seconds in the blue and green wires, respectively. A fast Fourier transform (FFT) of the
Figure 4.5: Results of single wire experiment in air with blue wire driven at 50 Hz with the sensor placed on the plane between the two wires. The blue curve represents the magnetic field strength of the first wire, \( H_1 \) (highlighted in blue in figure 4.4b), and the green curve represents that of second wire, \( H_2 \) (highlighted in green in figure 4.4b). Note that the green curve is zero here because the second wire was not being driven during this experiment. The amplitude of \( H_1 \) is 3.35 A/cm. The orange curve represents the strength of the magnetic field measured by the Hall effect sensors (\( xy \)-position highlighted in orange with black border in figure 4.4b). Each row corresponds to a measurement made by one of the three sensors. On the right are depicted the fast Fourier transform (FFT) magnitudes of the magnetic fields. Each FFT shows a peak at 50 Hz and is zero elsewhere, showing that the frequency of the current inputted matches that of the induced magnetic flux. The magnitudes of these peaks are 0.38 mT, 0.06 mT, and 0.16 mT for \( B_1 \), \( B_2 \), and \( B_3 \), respectively; and the calculated overall magnitude is 0.42 mT. This yields a permeability of \( \mu = \frac{B}{H} = \frac{0.42 \text{ mT}}{3.35 \text{ A/cm}} = 1.25 \times 10^{-6} \).
Figure 4.6: Results of the same experiment but with green wire driven at 50 Hz. Here, the blue curve is zero because only the green wire is being driven. The amplitude of the magnetic field strength $H_2 = 1.80 \text{ A/cm}$ is about 0.54 times that of the previous experiment (figure 4.5). The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.19 mT, 0.07 mT, and 0.05 mT, respectively; and the calculated overall magnitude is 0.21 mT. Elsewhere, the FFT is zero. The calculated permeability is $\mu = \frac{B}{H} = \frac{0.21 \times 10^{-3} \text{ T}}{180 \text{ A/m}} = 1.17 \times 10^{-6} \text{ H/m}$.

Induced magnetic field was calculated and the power (amplitude) of each of the peaks is reported alongside a 100 ms snapshot of the time series for each of the drive currents and the induced field. The result of driving only the blue wire at 50 Hz is shown in figure 4.5. We see that the frequency of the measured magnetic field is also 50 Hz along all three axes, which matches expectations. To verify that the magnetic field strength measurement was accurate, we calculated the permeability $\mu$, which is the ratio of induced magnetic field magnitude to magnetic field strength, i.e.:

$$\mu = \frac{B}{H} \quad (4.2)$$
Figure 4.7: Results of two wire experiment performed in air with both wires driven at 50 Hz with the same magnetic field strengths. Note that both the blue and green H curves are in phase and have approximately the same amplitude as before because they were driven by the same current. The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.18 mT, 0.01 mT, and 0.22 mT, respectively; elsewhere, the FFT is zero. This shows that the magnetic fields do in fact sum linearly in air ($|−0.38+0.19| = |−0.17| ≈ 0.17$ mT $≈ 0.18$ mT, $|0.06−0.06| = 0.00$ mT $≈ 0.01$ mT, and $|0.16 + 0.05| = 0.21$ mT $≈ 0.22$ mT, where the signs of the fields are flipped if the phase of $B$ with respect to $H$ is 180 degrees).

From figure 4.5, we see that the amplitude of the measured magnetic field is $|B_{\text{max}}| = \sqrt{B_1^2 + B_2^2 + B_3^2} = \sqrt{0.38^2 + 0.05^2 + 0.16^2} = 0.42$ mT, and the amplitude of the calculated applied magnetic field strength is 3.15 A/cm. Thus, $\mu = \frac{B}{H} = \frac{0.42 \times 10^{-3}}{315} \frac{T}{A/m} = 1.32 \times 10^{-6}$ H/m $≈ 1.257 \times 10^{-6}$ H/m $= \mu_{\text{air}}$ (Inan, 1998). Thus, we approximately recover the permeability of air from the ratio of $B$ to $H$, indicating that the calculation of magnetic field strength is valid. Error may be attributed to the deviation in the measurement of distance from the sensors to the wire.
Figure 4.8: Results of single wire experiment performed in air with green wire driven at 93 Hz. The blue curve is zero because only the green wire is being driven. The amplitude of the magnetic field strength $H_2$ is 1.76 A/cm. The magnitudes of the FFT peaks at 93 Hz for $B_1$, $B_2$, and $B_3$ are 0.19 mT, 0.06 mT, and 0.05 mT, respectively; and the calculated overall magnitude is 0.21 mT. Elsewhere, the FFT is zero. Here, we find the permeability to be $\mu = \frac{B}{H} = \sqrt{0.19^2 + 0.06^2 + 0.05^2} \text{ mT} = 1.76 \text{ A/cm} = 1.18 \times 10^{-6} \text{ H/m}$.

We now summarize the results of the experiments in air. For the single wire experiments in which only the first wire was driven, the frequency of the input matches that of the measurement and the two are either in phase or 180 degrees out of phase depending on whether the sensor was oriented “towards” or “away from” the wire. As stated above, the amplitude of the induced magnetic field is 0.42 mT for a magnetic field strength of about 3.15 A/cm for $H_1$. We then repeated the experiment but instead of passing current into the first wire, we drove the second wire at an amplitude of 18.1 A. This resulted in a magnetic field strength of about 1.70 A/cm. Figure 4.6 shows the result of measuring the magnetic field with the
sensors oriented as before. Notice that while the spectrum of the magnetic field is only nonzero at 50 Hz, the height of the peaks have decreased in proportion to the decrease in magnetic field strength. To determine whether the induced fields sum linearly in air, we measured the field at the same location when both wires were being driven by a 50 Hz signal at their respective current amplitudes. The result is given in figure 4.7. As expected by Maxwell’s equations, the magnetic fields induced by each of the wires sum up linearly, so that the strength of the induced field is the sum of the fields induced by each wire. Finally, we repeated the experiment with two different frequencies to determine whether there was a frequency-dependent nonlinear effect.
Figure 4.10: Example of skin effect in a conductor. Current $I$ flowing from bottom to top through the conductor induces a curling magnetic field $\mathbf{H}$. This curling magnetic field produces eddy currents $I_w$ of its own which oppose the flux generating it. The eddy currents $I_w$ negate the effect of the current $I$ in the center of the conductor but allow currents to flow along its surface.

Figure 4.8 first shows the result of driving the green wire with a frequency of 93 Hz and a magnetic field strength of 1.66 A/cm. Again, the spectrum of the magnetic field is only nonzero at 93 Hz. Finally, figure 4.9 shows the result of the experiment when the blue wire is driven by 50 Hz and the green wire is driven by 93 Hz at their respective amplitudes. As predicted by Maxwell’s equations, the magnetic fields sum linearly even when different frequencies of current are applied to the two wires.

In the next series of experiments, we explored different media to determine whether changing the medium in which the magnetic fields were induced would cause nonlinear effects. We started with 0.9% saline, which is a conductive solution that allows eddy currents to flow. These are generated to counteract a changing magnetic flux by producing a flux in the opposite direction, a principle known as Lenz’s law. In wire, eddy currents oppose the flow of current through the center of the wire which is a phenomenon known as the skin effect (Inan, 1998) (see figure...
Figure 4.11: Results of single wire experiment in saline with blue wire driven at 50 Hz with the sensor placed on the plane between the two wires. The blue curve represents the magnetic field strength of the first wire, $H_1$ (highlighted in blue in figure 4.4b), and the green curve represents that of second wire, $H_2$ (highlighted in green in figure 4.4b). Note that the green curve is zero here because the second wire was not being driven during this experiment. The amplitude of $H_1$ is 3.34 A/cm. The orange curve represents the strength of the magnetic field measured by the Hall effect sensors (xy-position highlighted in orange with black border in figure 4.4b). Each row corresponds to a measurement made by one of the three sensors. On the right are depicted the fast Fourier transform (FFT) magnitudes of the magnetic fields. Each FFT shows a peak at 50 Hz and is zero elsewhere, showing that the frequency of the current inputted matches that of the induced magnetic flux. The magnitudes of these peaks are 0.33 mT, 0.07 mT, and 0.15 mT for $B_1$, $B_2$, and $B_3$, respectively; and the calculated overall magnitude is 0.36 mT. This yields a permeability of $\mu = \frac{B}{H} = \frac{0.36 \text{ mT}}{3.34 \text{ A/cm}} = 1.09 \times 10^{-6}$. 

88
Figure 4.12: Results of the same experiment but with green wire driven at 50 Hz. Here, the blue curve is zero because only the green wire is being driven. The amplitude of the magnetic field strength $H_2 = 1.81 \, \text{A/cm}$ is about 0.54 times that of the previous experiment (figure 4.11). The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.22 mT, 0.08 mT, and 0.06 mT, respectively; and the calculated overall magnitude is 0.24 mT. Elsewhere, the FFT is zero. The calculated permeability is $\mu = \frac{B}{H} = \frac{0.24 \times 10^{-3}}{1.81} \, \text{T} = 1.35 \times 10^{-6} \, \text{H/m} \approx 1.26 \times 10^{-6} \, \text{H/m} = \mu_{\text{air}}$.

The higher the frequency of a signal, the greater the skin effect and the less volume that current has to flow within a conductor, increasing the resistance of the wire. The results of repeating the experiments performed above in air are shown in figures 4.11 through 4.15. In summary, we find that magnetic fields interact linearly in saline just as they do in air and that at a magnetic field strength $H$ on the order of $1 - 3 \, \text{A/cm}$ is insufficient to induce eddy currents strong enough in saline to generate nonlinear effects.

Finally, we repeated the above experiments in ferrofluid. Ferrofluid is a suspension of 3-15% iron oxide (magnetite) nanoparticles, 6-30% oil soluble dispersant, and
Figure 4.13: Results of two wire experiment performed in saline with both wires driven at 50 Hz with the same magnetic field strengths. Note that both the blue and green $H$ curves are in phase and have approximately the same amplitude because they are driven by approximately the same current. The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.10 mT, 0.01 mT, and 0.20 mT, respectively; elsewhere, the FFT is zero. This shows that the magnetic fields do in fact sum linearly in saline ($|−0.33+0.22| = |−0.11| = 0.11 \text{ mT} \approx 0.10 \text{ mT}$, $|0.07−0.08| = |−0.01| = 0.01 \text{ mT}$, and $|0.15 + 0.06| = 0.21 \text{ mT} \approx 0.20 \text{ mT}$, where the signs of the fields are flipped if the phase of $B$ with respect to $H$ is 180 degrees).

55-91% hydrotreated light distillates (petroleum) (Ferrotec, 2012). It was invented in 1963 by NASA as a type of rocket fuel whose injection could be controlled by a magnetic field. The properties of the nanoparticles allow ferrofluid to overcome magnetic attraction and the van der Waals force prevents it from clumping. This allows ferrofluid to dynamically conform to the orientation of magnetic field lines in its vicinity. Ferrofluid is thus also a useful tool for visualizing magnetic fields. Recall that magnetite is also the nanoparticle known to exist both in safe and pathological forms in the brains of humans (Maher et al., 2016). Thus, in addition to the
Figure 4.14: Results of single wire experiment performed in saline with green wire driven at 93 Hz. The blue curve is zero because only the green wire is being driven. The amplitude of the magnetic field strength $H_2$ is 1.77 A/cm. The magnitudes of the FFT peaks at 93 Hz for $B_1$, $B_2$, and $B_3$ are 0.22 mT, 0.07 mT, and 0.05 mT, respectively; and the calculated overall magnitude is 0.24 mT. Elsewhere, the FFT is zero. Here, we find the permeability to be $\mu = \frac{B}{H} = \frac{\sqrt{0.22^2 + 0.07^2 + 0.05^2} \text{ mT}}{177 \text{ A/cm}} = 1.35 \times 10^{-6} \text{ H/m}$.

magnetic properties of ferrofluid that make it appealing, it is also makes our model more biologically realistic. Because the magnetization of iron saturates above a certain magnetic field strength, we hypothesized that induced magnetic fields may mix nonlinearly (Carvell et al., 2010). Figure 4.12 shows the result of performing the single and two wire experiments in ferrofluid. For a magnetic field strength $H$ on the order of $1-3 \text{ A/cm}$ employed in this experiment, magnetic fields only mix linearly in ferrofluid. This agrees with our expectations since according to the datasheet for our EFH-1 ferrofluid, we would need to generate a magnetic field strength of 7960 A/cm to achieve a saturation magnetization of about 44 mT. Our magnetic field strength
Figure 4.15: Results of two wire experiment performed in saline with the blue wire driven at 50 Hz and the green wire driven at 93 Hz. The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.33 mT, 0.07 mT, and 0.15 mT, respectively; and those at 93 Hz for $B_1$, $B_2$, and $B_3$ are 0.22 mT, 0.07 mT, and 0.05 mT, respectively. Elsewhere, the FFT is zero. These match the magnitudes shown in figure 4.11 and 4.14 for the respect frequencies, showing that the interaction at different frequencies is also linear.

is about three orders of magnitude weaker, and the magnetic field induced is about two orders of magnitude less. Thus, it is likely that at this magnetic field strength, we are still operating within the linear range of ferrofluid.

In conclusion, while the wire model was a good first pass which mimics the configuration of the white matter pathways in the brain, it presents several disadvantages. First, because the tracts consist of only a single wire each, the induced magnetic fields are too weak to interact over large distances. Since fields decay as $1/r$ with distance $r$ from the source, we had a choice between (1) building a device with heavier wires capable of sustaining large magnetic field strengths or (2) building a smaller device.
Figure 4.16: Results of single wire experiment in ferrofluid with blue wire driven at 50 Hz with the sensor placed on the plane between the two wires. The blue curve represents the magnetic field strength of the first wire, $H_1$ (highlighted in blue in figure 4.4), and the green curve represents that of second wire, $H_2$ (highlighted in green in figure 4.4b). Note that the green curve is zero here because the second wire was not being driven during this experiment. The amplitude of $H_1$ is 3.34 A/cm. The orange curve represents the strength of the magnetic field measured by the Hall effect sensors (xy-position highlighted in orange with black border in figure 4.4b). Each row corresponds to a measurement made by one of the three sensors. On the right are depicted the fast Fourier transform (FFT) magnitudes of the magnetic fields. Each FFT shows a peak at 50 Hz and is zero elsewhere, showing that the frequency of the current inputted matches that of the induced magnetic flux. The magnitudes of these peaks are 0.45 mT, 0.11 mT, and 0.30 mT for $B_1$, $B_2$, and $B_3$, respectively; and the calculated overall magnitude is 0.55 mT. This yields a permeability of $\mu = \frac{B}{H} = \frac{0.55 \text{ mT}}{3.34 \text{ A/cm}} = 1.66 \times 10^{-6}$. 
Figure 4.17: Results of the same experiment but with green wire driven at 50 Hz. Here, the blue curve is zero because only the green wire is being driven. The amplitude of the magnetic field strength $H_2 = 1.45 \text{ A/cm}$ is about 0.43 times that of the previous experiment (figure 4.16). The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.20 mT, 0.03 mT, and 0.04 mT, respectively; and the calculated overall magnitude is 0.21 mT. Elsewhere, the FFT is zero. The calculated permeability is $\mu = \frac{B}{H} = \frac{0.21 \times 10^{-3} \text{ T}}{145 \text{ A/m}} = 1.45 \times 10^{-6} \text{ H/m}$.

with smaller wires that allow fields to interact in a smaller volume, hence avoiding the need for stronger fields. While one could increase the current to increase the strength of the induced magnetic field, magnetic field strength scales only linearly with current. Thus, doubling or tripling the current would only result in the field multiplying by a factor of two or three. Our audio amplifiers are limited to generating 700 W of continuous power. The maximum current deliverable by the amplifier would then be $I = \sqrt{\frac{P}{R}} = \sqrt{700/0.2} = 59 \text{ A}$, so we approach the limit of the hardware quickly. Increasing current also causes more heating, so additional measures would have to be taken to cool the device if it were to be used for an extended
Figure 4.18: Results of two wire experiment performed in ferrofluid with both wires driven at 50 Hz. Note that both the blue and green $H$ curves are in phase and have approximately the same amplitude because they are driven by approximately the same current. The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.26 mT, 0.08 mT, and 0.33 mT, respectively; elsewhere, the FFT is zero. This shows that the magnetic fields do in fact sum linearly in ferrofluid ($| - 0.45 + 0.20 | = | - 0.25 | = 0.25 \text{ mT} \approx 0.26 mBox mT$, $| 0.11 - 0.03 | = | 0.08 | = 0.08 \text{ mT}$, and $| 0.30 + 0.04 | = 0.34 \text{ mT} \approx 0.33 \text{ mT}$, where the signs of the fields are flipped if the phase of $B$ with respect to $H$ is 180 degrees).

period of time. On the other hand, shrinking down the device presents a problem since the magnetic fields would be too weak to detect by our Hall effect sensors, which are already near the limits of their sensing ability. More powerful equipment such as a SQUID or atomic magnetometer may provide a possible solution, but the former is expensive and the latter is not yet commercially available. Second, at the field strengths tested, neither the saline solution nor the ferrofluid resulted in nonlinear interactions of magnetic fields. As a result, computations that involve nonlinear transformations could not be performed with magnetic fields. Lastly, the model is
Figure 4.19: Results of single wire experiment performed in ferrofluid with green wire driven at 93 Hz. The blue curve is zero because only the green wire is being driven. The amplitude of the magnetic field strength $H_2$ is 1.41 A/cm. The magnitudes of the FFT peaks at 93 Hz for $B_1$, $B_2$, and $B_3$ are 0.20 mT, 0.03 mT, and 0.04 mT, respectively; and the calculated overall magnitude is 0.20 mT. Elsewhere, the FFT is zero. Here, we find the permeability to be $\mu = \frac{B}{H} = \frac{\sqrt{0.20^2+0.03^2+0.04^2} \text{ mT}}{1.41 \text{ A/cm}} = 1.43 \times 10^{-6} \text{ H/m}$.

only a very rough approximation of the geometry of the tracts in the brain.

4.4.2 Design 2: 3D printout

To address the fact that in the previous design, the tracts themselves were not realistic in shape (i.e. they did not realistically model the shape of the white matter tractography), we took a 3D printing approach to modeling the tracts. We hypothesized that by utilizing the full posterior distribution of tract samples outputted by TRACULA, we could come up with a very accurate trajectory along which a tract should grow. The final device we aimed to construct could reflect the statistics of
Figure 4.20: Results of two wire experiment performed in ferrofluid with the blue wire driven at 50 Hz and the green wire driven at 93 Hz. The magnitudes of the FFT peaks at 50 Hz for $B_1$, $B_2$, and $B_3$ are 0.45 mT, 0.11 mT, and 0.30 mT, respectively; and those at 93 Hz for $B_1$, $B_2$, and $B_3$ are 0.20 mT, 0.03 mT, and 0.04 mT, respectively. Elsewhere, the FFT is zero. These match the magnitudes shown in figure 4.5 and 4.8 for the respective frequencies, showing that the interaction of different frequencies is also linear.

The pathway distributions obtained by TRACULA in the form of 3D printed tubular tracts. The tracts could be hollow so as to be filled with conductive gel mimicking the properties of white matter. Experiments could then be conducted in which we pass currents through the device with varying patterns to induce electromagnetic fields. Two designs for the device were considered. The first is a simple design, which accurately reflects the locations of the pathways. The radius of the pathways is isotropic and constant as a function of position along a given pathway. The second iteration factors in spatially varying covariance of the pathway distribution and describes the uncertainty in the configuration of the pathways along each tract. In the following
Figure 4.21: Monte Carlo distribution of pathways for the forceps major for the single, healthy subject.

sections, we describe these designs in greater detail.

Preprocessing

TRACULA outputs a set of 1500 sampled paths for each of the 18 tracts that are inferred. An example of this tract distribution is given in figure 4.21 for the forceps major, which passes through the corpus callosum to join the occipital lobes of both brain hemispheres. Each of these paths can contain a different number of samples, typically ranging from under 50 to close to 100. To standardize the data, we performed interpolation with cubic B-splines (?) to produce 500 samples for each pathway. We next identified “mean pathways,” which were computed by averaging over the 1500 pathways at each of the 500 locations along each tract. Figure 4.22 shows the mean pathway computed for each of the 18 tracts. Dashed lines show one standard deviation in the x-, y-, and z-directions.

Iteration 1: Organic tracts with circular cross-sections

The mean pathways define the skeleton of the tract configurations. In the first device iteration, we create circular tubes whose radii represent the average standard
deviation of the pathway distribution in the principal directions of diffusion along each tract. To compute this, we perform the following series of steps on each mean tract. First, we numerically compute the gradient of tract growth as a function of position along the interpolated mean tract. Mathematically, if \( \mathbf{p}_{ij} = [p_x, p_y, p_z]_{ij} \) represents the point \( j \) on tract \( i \), we define \( \mathbf{g}_{ij} = \left[ \frac{dp_x}{dx}, \frac{dp_y}{dy}, \frac{dp_z}{dz} \right]_{ij} = \left[ \dot{x}, \dot{y}, \dot{z} \right]_{ij} \), where the dot represents the spatial derivative.

Our end goal is to create a tubular structure. To achieve this, we first reduce the three dimensional point cloud of information we have for each location along a tract (due to the distribution of pathways provided by TRACULA) to two dimensions.
Figure 4.23a depicts point clouds that represent the Monte Carlo distribution of fiber pathways for 4 of the 500 interpolated indices. For each position, we use the coordinates of the position and the spatial gradient to define a plane. Figure 4.23b depicts planes along with the normal vectors for the featured samples. Figure 4.23c shows the point clouds projected onto these planes. We chose to project the data onto these planes because they are oriented perpendicular to the direction along which the tract is growing. This is in contrast to a plane that might be obtained, for example, via linear regression, which might not lie perpendicular to the direction of tract growth.

Next, we perform principal component analysis (PCA) on the projected data. While the ambient dimension of the data is three, because we have projected the data onto a plane, the data lies on a two-dimensional subspace, and we end up with only two non-zero eigenvalues. Figure 4.23d shows the eigenvectors of the covariance matrix in color overlaid on the true data for the same four positions along the forceps major tract. The unit eigenvectors are scaled by the square roots of their corresponding eigenvalues, which give the standard deviations of the data along the principal directions of spread. The simplest way then to create the tubular structure is to average these standard deviations at each position and then compute the median of these to define the radius of the tube. After performing this calculation, we imported the mean tracts into the 3D modeling program Solidworks and extruded tubular solids whose radii are defined as above.

Figures 4.24-4.26 show the current stage of the printed design. While this solid cannot be filled, future work will resolve this by extruding cuts (rather than solids) from a similar model but with thicker tracts. In other words, we will shell out the interior of an identical model with thicker tracts using the tract diameters calculated for this model. Thus, the statistics of the tracts will still be reflected, but we will be able to fill the model with conductive material to perform experiments.
Figure 4.23: Processing steps to produce tubular structures.
Figure 4.24: Orthographic views of simple model with circular cross-sections. Blue solids represent tracts while red solids represent scaffolding structures for the model.
Figure 4.25: Isometric view of the simple model (with circular cross-sections). Blue solids represent tracts while red solids represent scaffolding structures to keep the tracts in place.

Iteration 2: Elliptical cross-sections

Rather than reducing both eigenvalues for each coordinate along the tract down to just one number, we can retain the information to build a more realistic model. We start with the empirical covariance matrix for each projected point cloud, which corresponds to the maximum likelihood estimate for the covariance matrix of a multivariate gaussian distribution. From multivariate normal theory, we know that
for data centered at \( \mu \), \( p \times 100\% \) of the density lies in the set of \( \mathbf{x} \)'s defined by 
\[
(\mathbf{x} - \mu)^\top \Sigma^{-1} (\mathbf{x} - \mu) \leq \chi^2_k(p),
\]
where \( \Sigma \) is the covariance and \( \chi^2_k(p) \) is the inverse cumulative distribution function (inverse CDF) which is the chi-squared density with \( k \) degrees of freedom evaluated at \( p \). This is the equation of an ellipse whose major and minor axes are defined by the eigenvectors and eigenvalues of \( \Sigma \). The inverse CDF of \( x \sim p(x) \) simply tells us the value of \( x \) for which \( \int_{-\infty}^x p(x) \, dx = C \), where \( 0 \leq C \leq 1 \). Assuming the empirical covariance matrix is equal to the true covariance matrix (which is valid since \( 1500 = n \gg d = 2 \)), we can substitute the empirical covariance matrix in the formula to acquire the set of points that falls within a “confidence” ellipse for a given \( C \). Figure 4.23e depicts the contours of these ellipses such that \( C \) (as defined above) equals 0.4 (i.e. theoretically, 40\% of the points should fall within the ellipse).
If we produce such an ellipse for each point along the tract, we see that the radii of the ellipses as well as the orientations of the major and minor axes change as a function of position along the tracts. This provides a realistic representation of the volume inside which 100\% of the white matter lies. Figure 4.23f shows this as a cloud of points for the forceps major of the healthy subject. Solidworks’ ScanTo3D function can then be used to wrap a mesh around the point cloud, producing a surface. Figures 4.27b-4.27c show the Solidworks rendering of these surfaces. This surface can be further processed and converted into a solid, as shown in figure 4.27d, which can be printed or used in other stages of design.

Figures 4.28-4.30 show different views of this processing procedure applied to all 18 tracts. While all of tracts here are in mesh if not surface form, once post-processing is applied to add thickness, we will have hollow tubes which we can fill with conductive gel (just as described in section 4.4.2 for the simple model).

In the end, we did not use the 3D printed models for a few major reasons. First and foremost, printing hollow tracts is difficult with polylactic acid (PLA), a standard material used for 3D printing. Because the diameter of the tracts is so small, they tend to fill up with dust and solid particles even though they are designed to be printed hollow. Thus, filling the tracts becomes very difficult or impossible. Second, the magnetic fields generated by the 3D printed model would be on the order of those produced by the wire model described in the previous section. As we found, these are too weak to induce any nonlinear interactions. Finally, even if we drove the tracts with greater current, they run the risk of melting or catching fire from the heat.

4.4.3 Design 3: wire model with coils

The magnetic field strengths in the first wire model and the 3D printed model were directly proportional to the current passing through the wires / tracts. The induced
(a) Isometric view of point cloud in Solidworks.
(b) Isometric view of surfaces produced from ScanTo3D.
(c) Detailed view of the right “arm.”
(d) Solid produced from the surfaces.

Figure 4.27: Views of the surfaces and solid of the forceps major created from point clouds in Solidworks for the model with elliptical cross-sections.
Figure 4.28: Orthographic views of Solidworks designs for the realistic model with elliptical cross-sections. Each gray surface contains hundreds of individual faces outlined in black. These surfaces were fitted using Solidworks’ ScanTo3D. The darker regions are meshes created by joining together the elliptical cross-sections. For these tracts, more extensive post-processing is required to convert the meshes into surfaces. Once valid surfaces are constructed, they can be filled in to produce solids like that shown in figure 4.27d.
flux density was relatively weak and incapable of producing any nonlinear effects in the different media tested. A stronger magnetic field strength would be advantageous for two reasons. First, the magnetic field could permeate further throughout the volume of the medium allowing for more distant interactions. Second, the medium may respond differently to a stronger magnetic field. For example, stronger eddy currents in saline may allow for more cancellation of induced fields, and in ferrofluid, stronger field strengths may lead to saturation of the field generated by the nanoparticles. While a substantial increases in the current drive of the 10 gauge wires would result in overheating, winding wires in coils would amplify the induced magnetic field by a factor proportional to the number of turns per unit length (Inan, 1998). Thus, a
Figure 4.30: Additional views of the surfaces of the realistic model.
Figure 4.31: Design 3: wire model of neuromagnetic reactor with twelve coils. Each of the loops is arranged based on the discovered tractography (shown in figure 4.2). As before, analog signals can be generated on a computer, amplified, and fed into different coils. Magnetic fields are measured by a 3D array of Hall effect probes, each consisting of three sensors for measuring the $x$-, $y$-, and $z$-components of the magnetic field. The magnetic fields are recorded by another NI-DAQ board and post-processed on the computer.
Figure 4.32: The numeric and color code shown here indicates the tracts from figure 4.2 to which each of the coils in this reactor design correspond. While in the tractography these tracts do not actually form closed loops, we extrapolated the tracts into loops for this design with the hypothesis that they might generate stronger magnetic fields that could interact nonlinearly in saline and ferrofluid media. Smaller loops of white matter may exist in the brain that behave similarly.

relatively small current could be applied to induce a high flux density.

To investigate this further, we built another wire model that consisted of coils of wire. Figure 4.31 shows this model, which is also about three times the scale of the human brain. The model is mounted inside an aluminum cage affixed with adjustable 3D printed holders for a set of 16 Hall effect probes. As before, each probe contains
three sensors measuring the flux density in three orthogonal directions. The design contains twelve coils each consisting of approximately 100 turns of 16 gauge magnet wire. Each coil corresponds to a different tract identified in the tractography study.
Figure 4.34: Result of coil experiment in air in which a 50 Hz current of amplitude 16.9 A was passed for four seconds through a coil of diameter 3 in. and consisting of 100 turns. The first column shows a 100 ms snapshot of the magnetic field strength and induced magnetic field. The magnetic field strength is shown in blue and has an amplitude of 179 A/cm. The magnetic flux densities measured by each of the three sensors is shown in orange and is either in phase or 180 degrees out of phase with the magnetic field strength depending on the orientation of the sensor. On the right side are depicted the FFTs of the measured magnetic field. Each FFT shows a peak at 50 Hz and is zero elsewhere, showing that the frequency of the current inputted matches that of the induced magnetic flux. The magnitudes of these peaks are 21.7 mT, 1.68 mT, and 2.97 mT for $B_1$, $B_2$, and $B_3$, respectively; and the calculated overall magnitude is 22.0 mT. $B$ and $H$ are about 2 orders of magnitude greater than the corresponding measurements for the single-wire experiment. From $B$ and $H$, we find the permeability to be $\mu = \frac{B}{H} = \frac{22.0 \text{ mT}}{179 \text{ A/cm}} = 1.23 \times 10^{-6} \text{ H/m}$.

(see figures 4.2 and 4.32). While in the tractography these tracts do not actually form closed loops, we extrapolated the tracts into loops for this design with the hypothesis that they might generate stronger magnetic fields that could interact nonlinearly in saline and ferrofluid media.
Figure 4.35: Result of coil experiment in air in which a 50 Hz current of increased amplitude 33.3 A was passed for four seconds through the same coil. The magnetic field strength increases to 354 A/cm. The induced magnetic fields increase to 42.8 mT, 3.47 mT, and 5.86 mT for each of the respective sensors, and the magnitude is 43.3 mT. The induced fields are nonzero only at 50 Hz. The ratio $\rho$ of magnetic field strengths for the 33.3 A and 16.9 A signals matches the ratio of induced magnetic fluxes: $\rho = \frac{354 \text{ A/cm}}{179 \text{ A/cm}} = 1.98 \approx 1.97 = \frac{43.3 \text{ mT}}{22.0 \text{ mT}}$. Thus, magnetic fields generated in air by a single frequency of current change linearly as a function of magnetic field strength.

To test this hypothesis, we performed a series of experiments, similar to the ones performed on the two-wire setup described above. We first isolated one of the coils consisting of 100 turns with inner diameter 3 in., outer diameter 3.5 in., and height 1.1 in. Current at one or more frequencies could be passed into the coil and the induced flux density could be measured. Inside, we placed a 3 in. diameter container to hold 150 mL of different media including air, saline, and ferrofluid. Along the center axis of the container and on the central plane of the coil, we mounted a
Figure 4.36: Result of coil experiment in air in which a 93 Hz current of amplitude 14.2 A was passed for four seconds through the same coil. The magnetic field strength is 151 A/cm. The induced magnetic flux densities measured by the three sensors are nonzero only at 93 Hz and have amplitudes 18.3 mT, 1.40 mT, and 2.51 mT; the calculated overall magnitude is 18.5 mT. From $B$ and $H$, we find the permeability to be $\mu = \frac{B}{H} = \frac{18.5 \, \text{mT}}{151 \, \text{A/cm}} = 1.23 \times 10^{-6} \, \text{H/m}$.

Hall effect probe with three sensors. This is where the induced field is predicted by Maxwell’s equations to be strongest. This setup allowed us to gauge the magnetic field strength and the types of interactions that took place within a given medium. A diagram of this setup (not to scale) is given in figure 4.33a, and the circuit diagram is given in figure 4.33b.

As in the previous series of experiments, we first determined whether our setup was sound by verifying in air (1) that we could recover the permeability of air (hence demonstrating that our calculation of magnetic field strength was accurate) and (2) that the frequencies of the inputted current and measured magnetic field matched.
Figure 4.37: Result of the coil experiment in air in which both 50 and 93 Hz sinusoids were inputted into the coil for four seconds. The amplitudes of the respective currents were 16.5 A and 14.0 A, similar to those in the above experiments. The corresponding magnetic field strengths are $H_1 = 175 \text{ A/cm}$ and $H_2 = 148 \text{ A/cm}$. We find that the magnetic flux density has nonzero frequency components at 50 Hz and 93 Hz. The amplitude of the 50 Hz components on the three sensors are 21.2 mT, 1.70 mT, and 2.89 mT; the amplitude of the 93 Hz components are 17.9 mT, 1.43 mT, and 2.43 mT. The 50 Hz component has magnitude 21.5 mT, and the 93 Hz component has magnitude 18.1 mT. The changes in flux densities at the two frequencies are proportional to the change in amplitude of the corresponding field strengths: $21.5 \text{ mT} = \left(\frac{175 \text{ A/cm}}{179 \text{ A/cm}}\right) \left(22.0 \text{ mT}\right) = 21.5 \text{ mT}$ and $18.1 \text{ mT} \approx \left(\frac{148 \text{ A/cm}}{151 \text{ A/cm}}\right) \left(18.5 \text{ mT}\right) = 18.2 \text{ mT}$. Thus, we find that, in air, magnetic fields induced by currents of different frequencies sum linearly.

Recall that permeability in H/m is the ratio of the induced magnetic flux density in T to the applied magnetic field strength in A/m. To determine the magnetic field strength of a solenoid whose length is much greater than its diameter, we calculate $H = nI$, where $n$ is the turns per unit length of the coil and $I$ is the current. (Hart,
Figure 4.38: Result of coil experiment in saline in which a 50 Hz current of amplitude 16.5 A was passed for four seconds through a coil of diameter 3 in. and consisting of 100 turns. The first column shows a 100 ms snapshot of the magnetic field strength and induced magnetic flux density. The magnetic field strength is shown in blue and has an amplitude of 175 A/cm. The magnetic flux densities measured by each of the three sensors is shown in orange and is either in phase or 180 degrees out of phase with the magnetic field strength depending on the orientation of the sensor. On the right side are depicted the FFTs of the measured flux densities. Each FFT shows a peak at 50 Hz and is zero elsewhere, showing that the frequency of the current inputted matches that of the induced magnetic flux. The magnitudes of these peaks are 21.1 mT, 1.86 mT, and 2.95 mT for $B_1$, $B_2$, and $B_3$, respectively; and the calculated overall magnitude is 21.4 mT. $B$ and $H$ are about 2 orders of magnitude greater than the corresponding measurements for the single-wire experiment. From $B$ and $H$, we find the permeability to be $\mu = \frac{B}{H} = \frac{21.4 \text{ mT}}{175 \text{ A/cm}} = 1.22 \times 10^{-6} \text{ H/m}$.

However, because our coil is short (1.1 in. tall) relative to the diameter (3 in.), we need to modify our formula; the formula for the magnetic flux density $B$ along the axis of a short, thick solenoid is given by (Hart, 2018):

$$\cos \beta = \frac{(x + L/2)}{[(x + L/2)^2 + R^2]^{0.5}}$$
Figure 4.39: Result of coil experiment in saline in which a 50 Hz current of amplitude 32.71 A was passed for four seconds through the same coil. The magnetic field strength increases to 347 A/cm. The induced magnetic flux densities increase to 41.9 mT, 3.73 mT, and 5.84 mT for each of the respective sensors, and the magnitude is 42.5 mT. The induced fields are nonzero only at 50 Hz. The ratio $\rho$ of magnetic field strengths for the 32.7 A and 16.5 A signals closely matches the ratio of induced magnetic fluxes: $\rho = \frac{347 \text{ A/cm}}{175 \text{ A/cm}} = 1.98 \approx 1.99 = \frac{42.5 \text{ mT}}{21.4 \text{ mT}}$. Thus, magnetic fields generated in air by a single frequency of current change linearly as a function of magnetic field strength.

$$\cos \alpha = \frac{(x - L/2)/[(x - L/2)^2 + R^2]^{0.5}}$$

$$B = \mu NI(\cos \beta - \cos \alpha)/(2L)$$

Here, $x$ is the position along the axis of the coil at which we would like to measure the flux density, $L$ is the height of the coil, $R$ is its radius, $N$ is the total number of turns, $I$ is the current, and $B$ is the resultant flux density. At a position of $x = 0$, the plane of the solenoid, we find that $\cos \alpha = - \cos \beta$, so the resultant flux density
Figure 4.40: Result of coil experiment in saline in which a 93 Hz current of amplitude 14.0 A was passed for four seconds through the same coil. The magnetic field strength is 148 A/cm. The induced magnetic fields measured by the three sensors are nonzero only at 93 Hz and have amplitudes 17.8 mT, 1.57 mT, and 2.48 mT; the calculated overall magnitude is 18.1 mT. From $B$ and $H$, we find the permeability to be $\mu = \frac{B}{H} = \frac{18.1 \text{ mT}}{148 \text{ A/cm}} = 1.22 \times 10^{-6} \text{ H/m}$.

is given by:

$$B = \mu H = \mu nI \cos \beta$$ \hspace{1cm} (4.3)

Using this formula, we calculated the magnetic field strength $H = nI \cos \beta$ for the first coil experiment, which was performed in air. In this experiment, the coil was driven by a 50 Hz sine wave of amplitude 16.9 A. The result of this experiment is given in figure 4.34. We find that the amplitude of the magnetic field strength is $H = nI \cos \beta = (100/(2.79 \text{ cm})) (16.9 \text{ A})(0.297) = 179 \text{ A/cm}$ and the magnitude of the induced magnetic field is $|B_{max}| = \sqrt{B_1^2 + B_2^2 + B_3^2} = \sqrt{21.7^2 + 1.68^2 + 2.97^2} = 22.0 \text{ mT}$. The resulting $\mu$ for air is $\mu = B/H = (22.0 \text{ mT})/(179 \text{ A/cm}) = 1.23 \times$
Figure 4.41: Result of the coil experiment in saline in which both 50 and 93 Hz sinusoids were inputted into the coil for four seconds. The amplitudes of the respective currents were 16.2 A and 13.8 A, similar to those in the above experiments. The corresponding magnetic field strengths are $H_1 = 172 \text{ A/cm}$ and $H_2 = 146 \text{ A/cm}$. We find that the magnetic field has nonzero frequency components at 50 Hz and 93 Hz. The amplitude of the 50 Hz components on the three sensors are 20.8 mT, 1.84 mT, and 2.88 mT; the amplitude of the 93 Hz components are 17.6 mT, 1.55 mT, and 2.44 mT. The 50 Hz component has magnitude 21.1 mT, and the 93 Hz component has magnitude 17.8 mT. The changes in flux densities at the two frequencies are proportional to the change in amplitude of the corresponding field strengths: $21.1 \text{ mT} \approx \left( \frac{172 \text{ A/cm}}{175 \text{ A/cm}} \right) (21.4 \text{ mT}) = 21.0 \text{ mT}$ and $17.8 \text{ mT} \approx \left( \frac{146 \text{ A/cm}}{148 \text{ A/cm}} \right) (18.1 \text{ mT}) = 17.9 \text{ mT}$. Thus, we find that, in saline, magnetic fields induced by currents of different frequencies sum linearly.

$10^{-6} \text{ H/m}$ which closely matches the magnetic permeability $\mu_{\text{air}} = 1.26 \times 10^{-6}$ reported in literature (Inan, 1998). Also note from the FFT plots shown in the second column of figure 4.34 that for all three sensors, the measured flux density is nonzero only at 50 Hz. This verifies that the measurement setup is working as intended.
**Figure 4.42:** Amplitude spectrum of magnetic field strength when 50 Hz sinusoid of amplitude 16.3 A was applied to a coil inside which sat a ferrofluid medium. Note the presence of harmonics of 50 Hz, particularly odd harmonics. These are induced as back emf in the coil due to as ferrofluid approaches magnetic saturation. Also present are weak harmonics at multiples of 60 Hz which are due to the power line. Despite efforts to shield the experiment from 60 Hz noise, these signals could not be completely blocked. Even harmonics are also present but at relatively weaker amplitudes.

Note that the value of permeability found above is much more accurate than the value found from the wire experiments. This is because the magnetic field produced by current in a single wire is much weaker than that produced by current in a coil. The magnetic field sensors have a resolution of 15 mV/mT, and the USB-6225 has an ADC with a resolution of 0.15 mV/bit. Thus, the flux density of the wire, which falls approximately in the range of ±1 mT, is represented by about \((1 \text{ mT})(0.015 \frac{V}{mT})\left(\frac{2^{16} \text{ levels}}{10 \text{ V}}\right) \approx 99\) levels discrete levels, and being off by just one level would introduce an error of over 1%. On the other hand, signals measured in the coil experiment fall approximately in the range of ±25 mT and are therefore represented by about \((25 \text{ mT})(0.015 \frac{V}{mT})\left(\frac{2^{16} \text{ levels}}{10 \text{ V}}\right) \approx 2458\) levels. Thus, an error
Figure 4.43: Result of coil experiment in ferrofluid in which a 50 Hz current of amplitude 16.3 A was passed for four seconds through a coil of diameter 3 in. and consisting of 100 turns. The first column shows a 100 ms snapshot of the magnetic field strength and induced magnetic field. The magnetic field strength is shown in blue and has an amplitude of 173 A/cm. The magnetic fields measured by each of the three sensors is shown in orange and is either in phase or 180 degrees out of phase with the magnetic field strength depending on the orientation of the sensor. On the right side are depicted the FFTs of the measured magnetic field. Unlike any of the previous results observed so far, we also observe magnetic fields of nonzero magnitude at all odd multiples of 50 Hz, such as 150 Hz, 250 Hz, and 350 Hz. These magnitudes decay exponentially and are listed in table 4.1. The presence of these odd harmonics suggests that the magnetic field induced in ferrofluid saturates above a certain magnetic field strength. Thus, magnetic fields generated in ferrofluid by a single frequency of current change nonlinearly as a function of magnetic field strength. $B$ and $H$ are about 2 orders of magnitude greater than the corresponding measurements for the single-wire experiment.

of 1 level would only produce an error of 0.04%. The greater amplitudes result in a higher dynamic range - $20 \log_{10} \left( \frac{25 \text{ mT}}{(15 \text{ mV/bit})(1 \text{ mT}/15 \text{ mV/bit})} \right) \approx 68 \text{ dB}$ compared to $20 \log_{10} \left( \frac{1 \text{ mT}}{(15 \text{ mV/bit})(1 \text{ mT}/15 \text{ mV/bit})} \right) \approx 40 \text{ dB}$ - which makes the
Figure 4.44: Hypothetical magnetic flux density induced by a sinusoidal field strength at 5 Hz. When the flux density saturates, the time series becomes a square wave. Plotting the spectrum, we find that a square wave consists of odd harmonics of the fundamental frequency (here 5 Hz) decaying geometrically.

We now proceed by discussing the results of the experiments. For the first experiment, in which the coil was driven by a 16.9 A current at 50 Hz, the magnetic field strength is 179 A/cm, and the amplitudes detected by each of the magnetic sensors at 50 Hz are 21.7 mT, 1.68 mT, and 2.97 mT, respectively (see figure 4.34). At other frequencies, the magnetic field induced is zero. The magnitude of 22.0 mT is about two orders greater than that induced by a single wire in section 4.4.1. Analytically, we find that for the single wire, the magnetic field strength for a current of 16.9 A at a distance of 1.5 in = 3.81 cm (the radius of the coil in this experiment) would be $H = \frac{I}{2\pi r} = \frac{16.9 \text{ A}}{2\pi(3.81 \text{ cm})} = 0.706 \text{ A/cm}$, which is a factor of 253 smaller than the magnetic field strength induced by this coil.
Figure 4.45: Amplitude spectrum of magnetic field strength when 50 Hz sinusoid of amplitude 32.6 A was applied to a coil inside which sat a ferrofluid medium. Note the presence of harmonics of 50 Hz, particularly odd harmonics. These are induced as back emf in the coil due to the magnetic saturation of ferrofluid.

Next, we compare these measurements to those generated by a current at the same frequency but at nearly double the amplitude. Figure 4.35 shows that when the current increased from about 16.9 A to about 33.3 A, the amplitude of the magnetic field strength as well as the measured flux density on all three sensors increases by about the same factor. The magnetic field strength at 50 Hz increases to 354 A/cm. The measured magnetic flux densities at 50 Hz are now 42.8 mT, 3.47 mT, and 5.86 mT, and the magnitude is 43.3 mT. Elsewhere, the magnetic field is zero. Thus, we find that $\rho = \frac{354 \text{ A/cm}}{179 \text{ A/cm}} = 1.98 \approx 1.97 = \frac{43.3 \text{ mT}}{22.0 \text{ mT}}$, where $\rho$ is the factor of increase. This indicates that, in air, the magnetic field at a single frequency is a linear function of magnetic field strength.

Finally, we generated sinusoids at both 50 and 93 Hz to determine whether nonlinear interactions may be frequency dependent. Figure 4.36 first shows the magnetic field strength and measured flux density for a 93 Hz, 14.2 A current inputted into the
Figure 4.46: Result of coil experiment in ferrofluid in which a 50 Hz current of amplitude 32.6 A was passed for four seconds through the same coil, emulating two individual sources at the same location. The magnetic field strength increases to 346 A/cm. We also observe magnetic fields of nonzero magnitude at all odd multiples of 50 Hz, such as 150 Hz, 250 Hz, and 350 Hz. These magnitudes decay exponentially and are listed in table 4.2.

The magnetic field strength is 151 A/cm at 93 Hz, and we find that again, the measured magnetic field is nonzero only at 93 Hz and is zero elsewhere. At this frequency, the amplitudes of the flux densities at the three sensors are 18.3 mT, 1.40 mT, and 2.51 mT; the calculated overall magnitude is 18.5 mT. Figure 4.37 shows the field strength and induced flux density for the coil when driven by currents at both 50 Hz and 93 Hz. The strength of the 50 Hz component is 175 A/cm, and that of the 93 Hz component is 148 A/cm. The measured flux density consists of nonzero components only at 50 Hz and 93 Hz. The 50 Hz components have amplitudes 21.2 mT, 1.70 mT, and 2.89 mT, and their magnitude is 21.5 mT; the 93 Hz components have amplitudes 17.9 mT, 1.43 mT, an 2.43 mT, and their magnitude is 18.1 mT. The changes in flux densities at the two frequencies are proportional to the change in amplitude of the
Figure 4.47: Result of coil experiment in ferrofluid in which a 93 Hz current of amplitude 13.2 A was passed for four seconds through the same coil. The magnetic field strength is 140 A/cm. The induced magnetic fields measured by the three sensors are nonzero only at 93 Hz and have amplitudes 13.2 mT, 0.580 mT, and 2.52 mT; the calculated overall magnitude is 13.4 mT.

Corresponding field strengths: 21.5 mT = \( \left( \frac{175 \text{ A/cm}}{179 \text{ A/cm}} \right) (22.0 \text{ mT}) \) = 21.5 mT and 18.1 mT \( \approx \left( \frac{148 \text{ A/cm}}{151 \text{ A/cm}} \right) (18.5 \text{ mT}) \) = 18.2 mT. Thus, we find that, in air, magnetic fields induced by currents of different frequencies sum linearly.

After showing that magnetic fields in air interacted linearly even at higher magnetic field strengths, we tested the hypothesis that stronger magnetic fields might interact nonlinearly in either saline or ferrofluid. We repeated the above experiments in these two media, starting with saline. Figure 4.38 shows the result of driving the same coil with a 175 A/cm magnetic field strength signal at 50 Hz. The measured magnetic flux density has a nonzero component only at 50 Hz. The amplitudes of the measured flux densities are 21.1 mT, 1.86 mT, and 2.95 mT for each of the three sensors, and
Figure 4.48: Result of the coil experiment in ferrofluid in which both 50 and 93 Hz sinusoids were inputted into the coil for four seconds. The amplitudes of the respective currents were 16.41 A and 13.26 A, similar to those in the above experiments. The corresponding magnetic field strengths are $H_1 = 170.9 \text{ A/cm}$ and $H_2 = 140.6 \text{ A/cm}$. We find that the magnetic field has nonzero frequency components at 50 Hz and 93 Hz. The amplitude of the 50 Hz components on the three sensors are 21.2 mT, 1.70 mT, and 2.9 mT; the amplitude of the 93 Hz components are 17.9 mT, 1.4 mT, and 2.4 mT. The 50 Hz component has magnitude 21.5 mT, and the 93 Hz component has magnitude 18.1 mT. The frequency components of the resulting magnetic fields are proportional to the change in amplitude: $21.5 \text{ mT} = \left( \frac{175.0 \text{ A/cm}}{175.9 \text{ A/cm}} \right) (21.95 \text{ mT}) = 21.5 \text{ mT}$ and $18.1 \text{ mT} \approx \left( \frac{148.0 \text{ A/cm}}{150.6 \text{ A/cm}} \right) (18.5 \text{ mT}) = 18.2 \text{ mT}$. Thus, we find that, in air, magnetic fields induced by currents of different frequencies sum linearly.

The overall magnitude is 21.4 mT. When the magnetic field strength is nearly doubled to an intensity of 347 A/cm, the induced flux densities at 50 Hz increase to 41.9 mT, 3.73 mT, and 5.84 mT, respectively, and the magnitude increases to 42.5 mT (see figure 4.39). This increase in amplitude is again proportional to the increase in magnetic field strength: $\rho = \frac{347 \text{ A/cm}}{175 \text{ A/cm}} = 1.98 \approx 1.99 = \frac{42.5 \text{ mT}}{21.4 \text{ mT}}$. Magnetic fields at all other frequencies have zero amplitude. Thus, increasing the magnetic field strength results
in a linear increase in induced magnetic field in saline. Next, a 93 Hz component with magnetic field strength 148 A/cm is generated, and flux densities of amplitude 17.8 mT, 1.57 mT, and 2.48 mT are measured at 93 Hz on the three sensors; the magnitude is 18.1 mT (see figure 4.40). When the 50 Hz and 93 Hz components are driven together with respective intensities 172 A/cm and 146 A/cm, we find that the induced flux densities at 50 Hz are 20.8 mT, 1.84 mT, and 2.88 mT (magnitude 21.1 mT) and at 93 Hz are 17.6 mT, 1.55 mT, and 2.44 mT (magnitude 17.8 mT) for each of the respective sensors (see figure 4.41). Magnetic fields at all other frequencies are zero. Once again, we see that the magnitude of the flux densities at each of the frequencies scales with magnetic field strength: 

\[
21.1 \text{ mT} \approx \left( \frac{172 \text{ A/cm}}{175 \text{ A/cm}} \right) (21.4 \text{ mT}) = 21.0 \text{ mT}
\]

and 

\[
17.8 \text{ mT} \approx \left( \frac{146 \text{ A/cm}}{148 \text{ A/cm}} \right) (18.1 \text{ mT}) = 17.9 \text{ mT},
\]

and therefore, magnetic fields induced by different frequencies interact linearly in saline.

We then repeated the experiments in ferrofluid, which, as was described in the previous section, contains magnetite nanoparticles whose magnetic flux reportedly saturates when the magnetic field strength exceeds 7960 A/cm (Ferrotec, 2018). So far, the largest magnetic field strength we have been able to achieve in our experiments is on the order of 350 A/cm at 50 Hz. This required a current of amplitude about 33.3 A. With 16 gauge wire, increasing the current further beyond this point and running the experiment for more than ten seconds at a time resulted in significant heating. Thus, the experiments we ran on ferrofluid were thought to be limited.

We first applied a current of frequency 50 Hz and amplitude 16.3 A to our coil and measured the induced magnetic flux density using the Hall effect probe. When measuring the magnetic field strength, we noticed that while a peak at 50 Hz of amplitude 173 A/cm was prominent, peaks also appeared at harmonics of this fre-
Table 4.1: Amplitudes of induced magnetic flux densities in ferrofluid at fundamental and harmonic frequencies due to 50 Hz input at 173 A/cm. The presence of odd harmonics suggests that the induced magnetic flux density is starting to saturate.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Frequency (Hz)</th>
<th>Amplitude (mT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>16.7</td>
</tr>
<tr>
<td>1</td>
<td>150</td>
<td>0.609</td>
</tr>
<tr>
<td>1</td>
<td>250</td>
<td>0.135</td>
</tr>
<tr>
<td>1</td>
<td>350</td>
<td>0.040</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>0.791</td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>0.078</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>0.011</td>
</tr>
<tr>
<td>2</td>
<td>350</td>
<td>0.002</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>3.14</td>
</tr>
<tr>
<td>3</td>
<td>150</td>
<td>0.057</td>
</tr>
<tr>
<td>3</td>
<td>250</td>
<td>0.016</td>
</tr>
<tr>
<td>3</td>
<td>350</td>
<td>0.005</td>
</tr>
</tbody>
</table>

Particularly prominent were the odd harmonics. An example of this is shown in figure 4.42. (Note that this spectrum as well as the remaining spectra reported for both field strength and flux density for ferrofluid are shown in log scale to highlight the presence of harmonics.) For example, at 150 Hz, we found an amplitude of 1.70 A/cm; at 250 Hz, 0.369 A/cm; and at 350 Hz, 0.113 A/cm. While these amplitudes are much weaker than that of the inputted signal, oscillations at these frequencies were not present in any of the previous experiments. Similarly, in the induced magnetic flux density, we noticed peaks at odd harmonics of 50 Hz. The amplitudes of these peaks are summarized in table 4.1. The time series and FFT plots are shown in figure 4.43. Even at a relatively low magnetic field strength relative to that reportedly required to magnetically saturate ferrofluid (Ferrotec, 2018), the material exhibits a nonlinear response at 50 Hz. The presence of dominant odd harmonics in the induced flux density suggests that the field is nearing saturation. When saturation is reached, the flux density plotted against time will take the shape of a square wave since after a certain point, its magnitude can no longer increase. A
Table 4.2: Amplitudes of induced magnetic flux densities in ferrofluid at fundamental and harmonic frequencies due to 50 Hz input with magnetic field strength 345.6 A/cm. The presence of odd harmonics suggests that the induced magnetic field is starting to saturate.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Frequency (Hz)</th>
<th>Amplitude (mT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>50</td>
<td>36.5</td>
</tr>
<tr>
<td>1</td>
<td>150</td>
<td>1.39</td>
</tr>
<tr>
<td>1</td>
<td>250</td>
<td>0.467</td>
</tr>
<tr>
<td>1</td>
<td>350</td>
<td>0.184</td>
</tr>
<tr>
<td>2</td>
<td>50</td>
<td>2.12</td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>0.244</td>
</tr>
<tr>
<td>2</td>
<td>250</td>
<td>0.059</td>
</tr>
<tr>
<td>2</td>
<td>350</td>
<td>0.018</td>
</tr>
<tr>
<td>3</td>
<td>50</td>
<td>6.26</td>
</tr>
<tr>
<td>3</td>
<td>150</td>
<td>0.102</td>
</tr>
<tr>
<td>3</td>
<td>250</td>
<td>0.046</td>
</tr>
<tr>
<td>3</td>
<td>350</td>
<td>0.021</td>
</tr>
</tbody>
</table>

Fourier series expansion of a square wave, which results in a sum of odd harmonic sine waves of geometrically decaying amplitude, thus explains the presence of odd harmonics (Haykin and Van Veen, 2007). This is depicted graphically in figure 4.44. Thus, there is a nonlinear relationship between applied field strength and induced flux density in ferrofluid.

We followed up with another experiment in which we applied a 50 Hz sinusoid of amplitude 32.6 A, double the amplitude used in the previous experiment. This generated a magnetic field strength of 346 A/cm at 50 Hz and also resulted in prominent odd harmonics being generated: 4.13 A/cm at 150 Hz, 1.33 A/cm at 250 Hz, and 0.519 A/cm at 350 Hz (see figure 4.45). The measured magnetic flux densities consisted of both fundamental frequency and (predominantly odd) harmonic components whose amplitudes are given in table 4.2. The time series and FFT plots are shown in figure 4.46. Note that the increase in magnetic flux density on each of the sensors is nonlinear and frequency dependent (i.e. a doubling of the magnetic field strength from 173 A/cm to 346 A/cm does not result in a doubling of the flux.
Table 4.3: Amplitudes of induced magnetic flux densities in ferrofluid at fundamental and harmonic frequencies due to 93 Hz input with magnetic field strength 140 A/cm. The presence of odd harmonics suggests that the induced magnetic flux density is starting to saturate.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Frequency (Hz)</th>
<th>Amplitude (mT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>93</td>
<td>13.2</td>
</tr>
<tr>
<td>1</td>
<td>279</td>
<td>0.466</td>
</tr>
<tr>
<td>1</td>
<td>465</td>
<td>0.085</td>
</tr>
<tr>
<td>1</td>
<td>651</td>
<td>0.024</td>
</tr>
<tr>
<td>2</td>
<td>93</td>
<td>0.580</td>
</tr>
<tr>
<td>2</td>
<td>279</td>
<td>0.054</td>
</tr>
<tr>
<td>2</td>
<td>465</td>
<td>0.005</td>
</tr>
<tr>
<td>2</td>
<td>651</td>
<td>0.000</td>
</tr>
<tr>
<td>3</td>
<td>93</td>
<td>2.52</td>
</tr>
<tr>
<td>3</td>
<td>279</td>
<td>0.048</td>
</tr>
<tr>
<td>3</td>
<td>465</td>
<td>0.011</td>
</tr>
<tr>
<td>3</td>
<td>651</td>
<td>0.003</td>
</tr>
</tbody>
</table>

density at each of the frequencies).

Next, we aimed to determine whether there was nonlinear interaction between magnetic fields induced at different frequencies. We first generated a 93 Hz sinusoid of amplitude 13.2 A. We measured the magnetic field strength and found that it again contained (predominantly odd) harmonics of 93 Hz; the field strength was 140 A/cm at 93 Hz, 1.40 A/cm at 279 Hz, 0.236 A/cm at 465 Hz, and 0.066 A/cm at 651 Hz. The induced magnetic flux density also contained (predominantly odd) harmonics, whose amplitudes are given in table 4.3 (time series and FFT plots given in figure 4.47). Then, we simultaneously inputted a 50 Hz signal of amplitude 16.1 A and a 93 Hz signal of amplitude 13.3 A. The magnetic field strength of this combination of signals as well as the induced magnetic flux density consists of harmonics of frequency $f_1$, $f_2$, $3f_1$, $3f_2$, $2f_1 \pm f_2$, and $2f_2 \pm f_1$, where $f_1 = 50$ and $f_2 = 93$. Furthermore, the harmonics themselves interact in the same way, producing even more frequency components. For instance, the second harmonic of 50 Hz (100 Hz) and the first harmonic of 93 Hz (93 Hz) interact to produce a 7 Hz oscillation $(2(50) - 93 = 7)$. 

131
This, in turn, interacts with the 93 Hz sinusoid to produce a 179 Hz oscillation \((2(93) - 7 = 179)\), and so on. The amplitudes of the prominent peaks are too many to list in table form, so we direct the reader to figure 4.48. We will explore the reasons for this effect further in section 4.4.4.

In conclusion, we have shown that for magnetic field strengths on the order of 100 to 300 A/cm, magnetic fields interact linearly in both air and saline; in ferrofluid, however, magnetic fields interact nonlinearly, producing a rich spectrum of sinusoids that consists of sums and differences of the applied base frequencies and their harmonics. In magnetic materials, the presence of two alternating magnetic fields is capable of generating harmonics of the form \(f_1, f_2, 3f_1, 3f_2, 2f_1 \pm f_2, \) and \(2f_2 \pm f_1\). The current design of the reactor allows magnetic fields to interact nonlinearly, but it is impractical to use for extended periods of time because the large current draw required to generate nonlinear magnetic fields of appreciable amplitude results in excessive heating. This is remedied in the final reactor design, which is discussed in section 4.4.5.

4.4.4 Nonlinear interaction of magnetic fields

So far, we have found that magnetic fields generated by alternating current sources interact linearly in air and saline and nonlinearly in ferrofluid. In this section, we explore in greater detail the cause of the nonlinear interaction of magnetic fields, verify it in a material called Nanoperm with a permeability much higher than ferrofluid (MAGNETEC GmbH, 2016), and explain how the effect can be used to generate a “soft” exclusive-OR (XOR) operation in analog. We first provide some background on the concept of a B-H curve, or hysteresis loop, which characterizes the induced magnetic flux density \((B, \text{units of T})\) as a function of the applied magnetic field strength \((H, \text{units of A/m or A/cm})\). The relationship between \(B\) and \(H\) is given
by the equation

$$B = \mu(H + M)$$ (4.4)

where $M$ is the magnetization of the material. Magnetization is a vector field that gives the magnetic moment at each location within a material (Chikazumi and Graham, 2009). In paramagnetic and diamagnetic materials $M$ aligns itself with or opposite to $H$; however, for ferromagnetic materials, the direction of $M$ can vary because they exhibit a property called remanence. Remanence, or residual flux density, is a property of ferromagnetic materials that quantifies how much flux density remains within the material after a previously applied magnetic field is removed. Soft ferromagnetic materials display a small amount of remanence and are commonly used in applications such as transformers in which magnetic losses and heating are points of concern while hard ferromagnetic materials have a large amount of remanence and are commonly used in hard disks, credit cards, or other settings in which permanent magnetization is beneficial. To demagnetize a ferromagnetic material, a magnetic field can be applied in the direction opposing its existing magnetization. The strength of this field is known as the coercive force. The dynamics of the magnetization of the material define the B-H curve, which shows how the magnetization changes as the applied magnetic field strength is varied.

Rabinovici and Kaplan state that for a conductive magnetic material with a nonlinear B-H curve, the relationship between $H$ and $B$ (the magnitudes of $H$ and $B$, respectively) can be modeled as (Rabinovici and Kaplan, 1983):

$$B = \mu(H + aH^3)$$ (4.5)

If the material is magnetized by two alternating currents of $f_1$ and $f_2$, the induced magnetic field will contain first harmonics $f_1$, $f_2$, $2f_1 \pm f_2$, $2f_2 \pm f_1$, $3f_1$, and $3f_2$. No magnetic field of low frequency $|f_1 - f_2|$, however, will be present. This can be seen by substituting for $H$ the sum of two sinusoids $\sin(2\pi f_1 t)$ and $\sin(2\pi f_2 t)$.
Expanding the cube and applying trigonometric identities shows that sinusoids at the above frequencies are present in $B$:

$$
B = \mu(H + H^3)
$$

$$
= \mu [(\sin(2\pi f_1 t) + \sin(2\pi f_2 t)) + (\sin(2\pi f_1 t) + \sin(2\pi f_2 t))^3]
$$

$$
= \frac{\mu}{4}[13 \sin(2\pi f_1 t) + 13 \sin(2\pi f_2 t) + 3 \sin(2\pi(2f_1 - f_2) t) + 3 \sin(2\pi(2f_2 - f_1) t) - 3 \sin(2\pi(2f_1 + f_2) t) - 3 \sin(2\pi(2f_2 + f_1) t) - \\
\sin(2\pi(3f_1 t) - \sin(2\pi(3f_2) t))]
$$

where we take $a = 1$ without loss of generality. Two conditions permit the presence of the low frequency signal: (1) the two currents are such that one is almost twice the frequency of the other or (2) the alternating currents are applied in the presence of a static magnetic field. In these cases, a square factor is added to the magnetic induction given in equation 4.5 and the additional magnetic field component is generated:

$$
B = \mu(H + H^3 + H^2)
$$

$$
= \mu [(\sin(2\pi f_1 t) + \sin(2\pi f_2 t)) + (\sin(2\pi f_1 t) + \sin(2\pi f_2 t))^3 + \\
\sin(2\pi f_1 t) + \sin(2\pi f_2 t))^2]
$$

$$
= \frac{\mu}{4}[13 \sin(2\pi f_1 t) + 13 \sin(2\pi f_2 t) + 3 \sin(2\pi(2f_1 - f_2) t) + 3 \sin(2\pi(2f_2 - f_1) t) - 3 \sin(2\pi(2f_1 + f_2) t) - 3 \sin(2\pi(2f_2 + f_1) t) - \\
\sin(2\pi(3f_1) t) - \sin(2\pi(3f_2) t) + 4 \cos(2\pi(f_1 - f_2) t) - 4 \cos(2\pi(f_1 + f_2) t) -
$$
\[2 \cos(2\pi(2f_1)t) - 2 \cos(2\pi(2f_2)t) + 4\] (4.6)

In solids, this magnetic field appears near the surface of but penetrates deep into the material, as opposed to the higher frequency fields which cannot penetrate as deeply due to cancellation by eddy currents (i.e. skin effect). Additional harmonics \(f_1 + f_2\), \(2f_1\), and \(2f_2\) are generated in addition to the \(f_1 - f_2\) component; moreover we see that a DC term also appears in equation 4.6, suggesting that the low frequency \(|f_1 - f_2|\) component can only be directly generated when a static magnetic field strength is applied.

In section 4.4.3, we showed the presence of harmonics when ferrofluid is magnetized. Here, we aim to verify the low frequency effect due to condition (2) in a high permeability, nanocrystalline alloy called Nanoperm (composition \(\text{Fe}_{73.5}\text{Cu}_{1}\text{Nb}_{3}\text{Si}_{15.5}\text{B}_{7}\)) (Total Materia, 2013). Nanoperm is a soft ferromagnetic conductor with a magnetic permeability of \(\mu = 80,000\) (MAGNETEC GmbH, 2016). The large permeability makes it easy to saturate with a relatively low magnetic field strength. Thus, relative to the currents that were employed to magnetize ferrofluid, the currents here are much weaker. We obtained toroidal tape-wound Nanoperm cores from Magnetec. Figure 4.49 shows the dimensions of the Nanoperm core side by side with a picture of one of the actual cores used in our device.

Figure 4.50a shows the B-H curve, or magnetic hysteresis loop provided by Magnetec. Macroscopic ferromagnetic materials (like the Nanoperm cores used here) contain domains in which magnetization is constant. This means that the magnetic dipoles of atoms within the domain are aligned. These domains are initially oriented randomly, minimizing the energy stored in the material, but when an external magnetic field is applied, they start to align themselves with the field. Eventually, all the domains align with the external field, and the magnetic flux density reaches a point called saturation. The higher the permeability of a material, the lower the amount
Figure 4.49: Nanoperm toroidal core. OD \( \leq 43.5 \) mm. ID \( \geq 22.5 \) mm. H \( \leq 18.5 \) mm (MAGNETEC GmbH, 2016). The four coils were used to drive two AC signals and one DC signal and to measure the induced magnetic field.

of magnetic field strength needed to saturate the magnetic flux density. According to figure 4.50a, a magnetic field strength of about 80 mA/cm is needed to saturate Nanoperm at a frequency of 50 Hz. (This is about \( 10^3 \) to \( 10^5 \) times less than the field strengths that were applied in the coil and wire models discussed previously.)

As the material saturates, the permeability relative to that of free space drops to 1, i.e. no more magnetic flux can be induced in the material because the magnetic domains are all already aligned.

Prior to conducting our experiments we again attempted to validate our setup by measuring the magnetic hysteresis of Nanoperm and comparing it to the curve reported in the datasheet (MAGNETEC GmbH, 2016). To do this, we wound two coils of 16 gauge magnet wire around the core. Because the permeability of Nanoperm is so high and a relatively small current can saturate the core, we wound only six turns for each coil. A circuit diagram of the measurement setup is given in figure 4.51.
Figure 4.50: Magnetic properties of Nanoperm shown in red. The magnetic flux density saturates at about 1.2 T for a magnetic field strength of about 80 mA/cm. Furthermore, the permeability of Nanoperm decreases as we increase the frequency of the input: after about 800 Hz, the permeability starts to significantly roll off, with a full-width half-maximum of about 12 kHz. Note that Nanoperm, like many soft ferromagnetic conductors, exhibits a hysteresis loop with a small coercive force.
Figure 4.51: Circuit used to measure hysteresis in Nanoperm core. A 200 mΩ shunt resistor is connected in series with the primary windings (1) to measure the current, which is proportional to the applied magnetic field strength, through the primary windings and (2) to ensure the coil does not overheat during measurement. The resistance is kept small to allow a significant voltage to drop across the coil itself. On the secondary side, we measure the voltage directly on the winding. By Faraday’s law, this voltage is proportional to the change in magnetic flux. Thus, integrating this voltage gives a quantity that is proportional to the flux itself. In our setup, integration was performed numerically using MATLAB.

The primary circuit consists of the primary coil, a 200 watt 200 mΩ shunt resistor (Electronics, 2007), and a Cerwin Vega CV-2800 audio amplifier connected in series (Cerwin-Vega!, 2011). The current through the primary winding and the voltage induced across the secondary winding are proportional to the applied magnetic field strength and change in magnetic flux density, respectively. Integrating the latter gives the magnetic flux density itself. These relationships are given by the following equations:

$$H \equiv \frac{V_R(t) \cdot N}{R \cdot l_c}$$  \hspace{1cm} (4.7)

where $V_R$ is the voltage across the shunt resistor, $N = 6$ is the number of turns, $R = 200 \text{ mΩ}$ is the resistance of the shunt resistor connected to the audio amplifier, and $l_c = 10.03 \text{ cm}$ is the magnetic path length given on the product datasheet.
where $E$ is the electromotive force (emf, units of volts) induced in the secondary winding, $N = 6$ is the number of turns, and $A_c = 0.88 \text{ cm}^2$ is the cross-sectional area of the core given in the datasheet (MAGNETEC GmbH, 2016). Trapezoidal numerical integration was performed in MATLAB. Figure 4.50b shows a plot of the measured hysteresis curve at 50 Hz in red, which closely matches the hysteresis loop given in the datasheet (see figure 4.50a). The flux density saturates at about $\pm 1.1$ T for an applied magnetic field strength of 90 mA/cm.

The permeability of ferromagnetic materials is also sensitive to frequency. Figures 4.50c and 4.50d show that at 10 kHz, Nanoperm saturates at a magnetic field strength of 90 mA/cm. Therefore, we can conclude that between 50 Hz and 10 kHz, as long as we are able to achieve a magnetic field strength of at least 90 mA/cm, we will be able to drive the magnetic flux into saturation, and thus into the nonlinear range of operation of the core.

Next, we conducted a series of experiments to verify the effect reported in (Rabinovici and Kaplan, 1983). Figure 4.52 shows the circuit diagram of the setup. Three primary windings were used to carry two alternating currents (ACs) and one direct current (DC) which induce alternating and direct magnetic fields in the core, respectively. One secondary winding was used to measure this induced field (via the same mechanism as was used to measure the hysteresis property of Nanoperm). We first left the DC loop open and inputted AC at frequencies of 50, 51, 67, 95, 102, 127, 147, and 150 Hz one at a time into one of the AC loops. These frequencies were randomly sampled and representative of the range 50 Hz to 150 Hz. Figure 4.53 shows the hysteresis loops for each of these cases, and figure 4.54 shows the amplitude spec-
Figure 4.52: Circuit used to verify nonlinear interaction of alternating magnetic fields in the presence of a direct magnetic field with a ferromagnetic conductor. The circuit consists of three primary windings and one secondary winding. Each winding has six turns. Two are connected to a shunt resistor of 200mΩ and are driven by AC; the other is driven by DC. No resistor is needed to measure the magnetic field strength for the winding with the DC source because the power supply outputs the current directly on its display (operated in current-limiting mode). To characterize the hysteresis at 50 Hz, the voltage induced in the secondary coil is numerically integrated.

The Fast Fourier Transform (FFT) of the magnetic field for the corresponding inputs. The hysteresis loops are practically identical, except that at 150 Hz, we start to see more noise due to the higher impedance of the core. (The impedance $Z_L$ of an inductor / transformer increases as a function of frequency: $Z_L = j\omega L$. This reduces the voltage applied across the shunt resistor which is used to measure the magnetic field strength, hence degrading the signal.) Notice, however, that the frequency content consists of not only the base frequency but also odd harmonics 3, 5, 7, etc. As before, this can be understood by performing a Fourier expansion of a square wave - the shape of the magnetic flux time series - which is known to consist only of odd harmonics (Haykin and Van Veen, 2007). Unlike for ferrofluid, we find that the even harmonics are not
Keeping the DC loop open, we then measured the hysteresis loop and the FFT of the induced field when pairs of these frequencies were inputted simultaneously. Figures 4.55 and 4.56 show the hysteresis loops and FFT of the magnetic field, respectively, for selected pairs. Notice that due to the saturation of the core, the induced magnetic fields combine nonlinearly, and we end up with many more frequency components in the measured field. As predicted by Rabinovici and Kaplan (1983), we see harmonics of $f_1$, $f_2$, $2f_1 - f_2$, $2f_2 - f_1$, $2f_1 + f_2$, $2f_2 + f_1$, $3f_1$, $3f_2$, and so on in figure 4.56, where $f_1$ and $f_2$ are the two input frequencies. However, a low frequency component $|f_1 - f_2|$ is not present.

Next, we measured the same hysteresis loops as above but in the presence of a direct magnetic field induced by passing DC into the second primary winding. The amount of current dictates the strength of the direct magnetic field. The power supply used to generate DC (BK Precision 1688B, (BK Precision, 2012)) provides a readout of the current. We repeated the experiment with four currents: 0.1 A, 0.5 A, 1 A, and 5 A, corresponding to magnetic field strengths of 59.8 mA/cm, 299.1 mA/cm, 598.2 mA/cm, and 2991.0 mA/cm, respectively (calculated via equation 4.7, where here, $V_R(t)/R$ is replaced by the current). For conciseness, in figures 4.57 and 4.58, we show hysteresis loops and FFTs of the induced magnetic field here for all four field strengths only for the 50 Hz signal; for all other frequencies, we give only the corresponding plots for a field strength of at 2991.0 mA/cm (for reasons stated below) in figures 4.61 and 4.62 and show the remaining plots in appendix A.

From figure 4.57, we see that as we increase the direct magnetic field strength $H_{DC}$, the hysteresis loop widens to the left. In other words, the coercive force needed to demagnetize the core to zero after positive magnetization increases in magnitude as the applied DC magnetic field strength increases. This trend is common to all frequencies (see appendix A) and is due to the fact that the external direct magnetic
Figure 4.53: Nanoperm hysteresis loops measured during single frequency inputs.
Figure 4.54: FFT of measured magnetic field for single frequency inputs. Note that the base frequency and the odd harmonics are both apparent in each figure. The amplitudes of the odd harmonics, fall off at approximately the same rate as that of a square wave, whose coefficients are inversely proportional to the order of the harmonic.
Figure 4.55: Nanoperm hysteresis loops measured when driven by two frequencies of AC without static magnetization.

Field magnetizes the material in one direction and either reinforces or opposes the alternating field. This also results in a short tail to the left and a long tail to the right of the hysteresis loop. When $H_{DC} = 59.8$ mA/cm, the negative coercive force is $H_{AC} = 1450$ mA/cm; on the other hand, when $H_{DC} = 2991.0$ mA/cm, the negative coercive force is $H_{AC} \approx 3950$ mA/cm.

In addition, we see that the scale of the $x$-axis has changed from tens of mA/cm (without DC, in figure 4.50b) to thousands of mA/cm, a trend that is also common to all frequencies (see appendix A). This happens because the direct current provides a sort of “magnetic inertia” which the alternating current must overcome. The static field produced by the DC source magnetizes the core in one direction, and the AC must fight against this to magnetize the core in the opposite direction. Next, when
Figure 4.56: FFT of magnetic field when core is driven by two frequencies without static magnetization. As predicted by Rabinovici and Kaplan (1983), we see harmonics of $f_1$, $f_2$, $2f_1 - f_2$, $2f_2 - f_1$, $2f_1 + f_2$, $2f_2 + f_1$, $3f_1$, $3f_2$, and so on, where $f_1$ and $f_2$ are the fundamental frequencies. However, a low frequency component $|f_1 - f_2|$ is not present.

The AC reverses polarity, the magnetized core creates a back emf in both the amplifier and the current source, resisting magnetization in the other direction. This process repeats over and over, forming the hysteresis loop.

In figure 4.58, we see that the FFT of the induced magnetic field for each of the direct magnetic field strengths. We see that for $H_{DC} = 59.8$ mA/cm, only the base frequency and its odd harmonics are present in the FFT; this is expected since the plot of the magnetic field as a function of time is essentially a square wave whose Fourier series representation consists of a sum of odd harmonics of a sinusoid. On the other hand, for the higher static magnetic field strengths, we see many additional
Figure 4.57: 50 Hz hysteresis loops measured in the presence of direct magnetic field at various field strengths. Note that for $H_{\text{DC}} = 299.1$ mA/cm and $H_{\text{DC}} = 598.2$ mA/cm, there seem to be two overlapping hysteresis loops. We suspect that the inner loop is an artifact but the reason for its presence is unclear.

frequency components which are not just harmonics of the base frequency. This result is tied to the hysteresis loop shifting leftward, as noted above. The shift indicates that there is greater magnetization in one direction than in the other. Unwrapping the hysteresis plot in figure 4.57d into a time series in figure 4.59, we find that the magnetic field spends more time in the positive magnetization than in the negative magnetization when the static magnetic field is on. While the general shape of the time series is still a square wave, it is no longer symmetric, resulting in the addition of frequency components. We further explore this in figure 4.60.

Finally, we show that for two alternating magnetic fields in the presence of a static magnetic field induced by passing DC in the third primary winding, we observe the
Figure 4.58: FFT of magnetic field when driven by 50 Hz at various direct magnetic field strengths.

Low frequency component at \( f^* = |f_1 - f_2| \). Figure 4.63 gives the hysteresis loops and figure 4.64 gives the FFT of the magnetic field. In comparison to the FFTs shown in figure 4.56, we find that the presence of a static magnetic field causes the difference frequency to appear as a peak. Interactions between \( f^* \) and the other harmonics results in more differences in the spectral pattern, and importantly, because the difference of any two frequencies can now be generated as a harmonic, many more lower frequency peaks are present. We note that the stronger the static magnetic field strength, the greater the amplitude of the difference frequency and associated harmonics. For the 59.8 mA/cm static field, almost no effect was observed. We believe that this is because even though at low values of DC, the core is able to saturate in one direction, the field strength required to reverse the saturation is on
Figure 4.59: 100 ms snapshot of the induced flux density due to 50 Hz sinusoidal input with and without the addition of a static magnetic field. We see that when a static magnetic field is present, the magnetization remains about 25% longer in the positive region than in the negative region while when no static magnetic field is present, the time spent in each region is about equal.

the order of thousands of mA/cm (as observed from figures 4.62 and 4.63). Thus, if the magnitude of the static field is too weak relative to the oscillating fields, the effect does not take place. Rabinovici and Kaplan also found that to elicit the effect, the static and alternating magnetic field strengths needed to be of similar magnitude (Rabinovici and Kaplan, 1983). We depict the hysteresis loops and FFT plots for the 2991.0 mA/cm field strength here and leave the corresponding plots at with experiments done at lower static magnetic field strengths for appendix A (figures A.8 through A.13). Furthermore, table 4.4 compares the magnitude of the flux density at $f^*$ for each of the four cases.

We now show that the low frequency output of a ferromagnetic conductor subjected to two alternating magnetic fields and a direct magnetic field resembles a smoothed, or “soft,” XOR function which we can use to perform nonlinear pattern
Figure 4.60: The top subplot depicts a simulated flux density which is positively saturated (value of 0.5 T) for 111 ms and negatively saturated (value of -0.5 T) for 89 ms. These are approximately the same durations as those found in figure 4.59. The bottom subplot shows its power spectrum, which consists of both even and odd harmonics in approximately the same ratios as found in figure 4.58d.

recognition. This operation is the cornerstone of the final design of the reactor. The truth table for the logical XOR operator is given in table 4.5. The logical XOR operator only operates on low (0) or high (1) inputs and only generates low or high outputs. Thus, the inputs and outputs are binary. When the inputs are the same the output is low; when they are different, the output is high. If we relax this property and allow the inputs and outputs to take on continuous values between 0 and 1, we obtain the absolute difference operation. A plot of the output of the absolute difference function on inputs ranging from 0 to 1 is given in figure 4.65a. We also find in figure 4.65b that by adjusting the range from [0, 1] to [50, 150], we experience no loss of generality.

From the above experiments on the Nanoperm cores, we know that in the presence of a static magnetic field, we can elicit a magnetic flux which oscillates at the absolute
Figure 4.61: Nanoperm hysteresis loops measured during single frequency inputs in the presence of 2991.0 mA/cm direct magnetic field strength.
Figure 4.62: FFT of measured magnetic field for single frequency inputs in the presence of 2991.0 mA/cm direct magnetic field strength.
\textbf{Figure 4.63:} Nanoperm hysteresis loops measured when driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 2991.0 mA/cm.

difference of two alternating fields. Thus, we should be able to perform an absolute difference operation using the cores by detecting the low frequency oscillation in the output. To test this, we inputted ACs for two seconds at magnetic field strengths between 2500 and 5000 mA/cm at all pairs of frequencies between 50 Hz and 99 Hz (one per coil), and measured the power of all frequencies between 0 and 49 Hz. We then repeated the experiment when the DC coil was also driven by a field strength of 2991.0 mA/cm. Note that we reduced the upper end of the range from 150 Hz to 99 Hz. This is because we want to restrict the difference frequency to fall outside the range of the minimum and maximum of the input frequencies; otherwise, there would be ambiguity as to whether a signal being measured is that of a difference frequency or an input frequency. For example, looking back at figure 4.64b, we would not be
Figure 4.64: FFT of magnetic field when core is driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 2991.0 mA/cm.

able to easily discern from the measurement whether the 96 Hz signal was an input or a harmonic that was generated. On the other hand, if inputs must lie between 50 Hz and 99 Hz, the difference frequencies must lie between 0 and 49 Hz.

Note that while the difference frequency is produced in the presence of a static magnetic field, its harmonics (2nd, 3rd, etc. as well as its interactions with harmonics of the inputs) are also generated. Many of these frequencies also lie in the range 0 to 49 Hz. Thus, picking the exact difference frequency from the FFT can be difficult. Instead, we resort to an approximation in which we take an average of the frequencies between 0 and 49 Hz weighted by their FFT powers. Mathematically, we perform the following to determine the weighted average difference frequency $\hat{f}^*$:

$$\hat{f}^* = \frac{\sum_i f_i P(f_i)}{\sum_i P(f_i)}$$  \hfill (4.9)
Table 4.4: Flux densities at difference frequencies with and without static magnetic field. Note that $B(|f_1 - f_2|)$ with DC is about two orders of magnitude greater than that without DC.

<table>
<thead>
<tr>
<th>$f_1$</th>
<th>50</th>
<th>51</th>
<th>95</th>
<th>127</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_2$</td>
<td>67</td>
<td>147</td>
<td>102</td>
<td>150</td>
</tr>
<tr>
<td>$B(f_1)$ without DC</td>
<td>929</td>
<td>1087</td>
<td>931</td>
<td>982</td>
</tr>
<tr>
<td>$B(f_2)$ without DC</td>
<td>829</td>
<td>695</td>
<td>860</td>
<td>802</td>
</tr>
<tr>
<td>$B(f_1)$ with DC</td>
<td>970</td>
<td>1164</td>
<td>862</td>
<td>888</td>
</tr>
<tr>
<td>$B(f_2)$ with DC</td>
<td>761</td>
<td>535</td>
<td>831</td>
<td>765</td>
</tr>
<tr>
<td>$B(</td>
<td>f_1 - f_2</td>
<td>)$ without DC</td>
<td>3.73</td>
<td>0.660</td>
</tr>
<tr>
<td>$B(</td>
<td>f_1 - f_2</td>
<td>)$ with DC</td>
<td>239</td>
<td>110</td>
</tr>
</tbody>
</table>

Table 4.5: Truth table for exclusive-OR (XOR) operator.

<table>
<thead>
<tr>
<th>$x$</th>
<th>$y$</th>
<th>out</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>

where $i$ indexes the frequencies of the discrete FFT performed in software, $f_i$ are these frequencies, and $P(f_i)$ are their powers in mT. We show that $\hat{f}$ comes closer to the true difference frequency when a static field is present compared to when one is not. For example, for ACs of frequency 53 Hz and 81 Hz and field strength approximately 4400 mA/cm, figure 4.66a shows a 250 ms snapshot of the time series and the FFT of the magnetic flux density when no static magnetic field is present, and figure 4.66b shows the same but in the presence of the static magnetic field.

By applying equation 4.9 to each of these cases we find that the weighted average frequency in the case without the static magnetic field is 19.6 Hz while the same for the case with the static magnetic field is 22.2 Hz. Thus the approximation to the true difference of 28 Hz is more accurate when DC is present compared to without.

The range of frequencies also affects the quality of the approximation. If the input frequencies are restricted to be too close, the dynamic range of the inputs suffers. On the other hand, if the input frequencies are too far apart, the approximation suffers
(a) Output of the absolute difference function \( z(x, y) = |x - y| \) for \( \{x, y\} \in [0, 1] \).

(b) Output of the absolute difference function \( z(x, y) = |x - y| \) for \( \{x, y\} \in [50, 150] \).

Figure 4.65: Theoretical absolute difference operations for two ranges. When the inputs \( x \) and \( y \) are similar in magnitude, the output is low; when they are different, the output is large.

(a) Without static magnetic field.

(b) With static magnetic field.

Figure 4.66: Each of the top subplots shows a 250 ms snapshot of the input magnetic field strength and the measured magnetic flux density for a 53 Hz and 81 Hz input. The ACs generated magnetic field strengths of 4510 mA/cm and 4370 mA/cm for the 53 Hz and 81 Hz components, respectively. The static field had strength 2991.0 mA/cm. The lower subplots show the FFT of the flux density (taken over the full 2 seconds of the experiment). We find that when the static magnetic field is present, the FFT shows a peak at 28 Hz which is absent when no static magnetic field is present.
because many harmonics are generated, washing out the low frequency effect of the static magnetic field. To optimize the range over which our approximation yielded the closest value to the truth while also providing a large enough dynamic range for possible input values, we simulated ranges from 2 Hz to 49 Hz. Figure 4.67 shows the correlation between the true difference frequency ($\Delta f$) and that predicted by the weighted average methods ($\hat{\Delta f}_1$ and $\hat{\Delta f}_2$ for the without and with DC models, respectively). The method with DC dominates over that without DC in all cases and achieves its highest correlations in the range $\Delta f \in [11, 24]$. Thus, we selected 20 to be the maximum difference in frequencies between the two inputs used in the reactor design, discussed in section 4.4.5.

In figure 4.68, we show the weighted average difference frequency predicted by each of the two models as well as the ground truth for all possible inputs in the
(a) Without static magnetic field.
(b) With static magnetic field.
(c) True differences.

Figure 4.68: Weighted average difference frequency predicted by each model and ground truth difference frequency. This figure confirms that predictions made by the model with the static magnetic field more closely match the ground truth difference frequency than those made by the model without the static field.

range [50, 70], a range with a maximum difference of 20 Hz. The picture confirms that the correlation between the predictions made by the model with DC is greater than those made by the model without DC.

4.4.5 Design 4: system of Nanoperm toroidal cores

Our final design of the reactor depicted in figure 4.69 utilizes the Nanoperm cores discussed in the previous section to perform nonlinear transformation on its inputs. The device is roughly the size of the human brain and consists of seven Nanoperm cores, which correspond to one or more tracts discovered in the tractography, as labeled in the figure. Each core has four coils of 16 gauge magnet wire wound around it with six turns each. The coils allow for the input of two ACs and one DC; the fourth coil is used to measure the induced flux density. The circuit diagram for each core is the same as that given in figure 4.52. Our design currently restricts each of these cores to function as a hierarchy of parallel computers. We reserve discussion of how these cores function together until the final chapter.
Figure 4.69: Design 4: neuromagnetic reactor made of toroidal Nanoperm cores. Each blue toroid contains a high permeability ($\mu_r = 80,000$) Nanoperm core that allows induced alternating magnetic fields to interact nonlinearly in the presence of a static magnetic field at relatively low amplitudes of current.
The neuromagnetic reactor described in the previous chapter constitutes the analog component of our hybrid computer. In this chapter, we describe our hardware perceptron, which comprises the digital component of the hybrid computer.

5.1 The perceptron learning algorithm

A perceptron is a model that classifies linearly separable data into one of two groups (Bishop, 2006). Inputs $x \in \mathbb{R}^D$ are linearly combined using a set of weights $w \in \mathbb{R}^D$, and the binary output $\hat{y} \in \{-1, 1\}$ is determined by a threshold function:

$$\hat{y} = \begin{cases} 1, & w^T x \geq t \\ -1, & w^T x < t \end{cases}$$

where $t$ is the threshold. The variable $x$ is a vector whose last component can be set to 1, allowing us to group the threshold parameter $t$ into the vector of weights $w$. In other words, by subtracting $t$ from both sides and appending a 1 to $x$, we force the
last entry of \( \mathbf{w} \) to be \(-t\). This results in a slightly simpler expression for prediction:

\[
\hat{y} = \begin{cases} 
1, & \mathbf{w}^\top \mathbf{x} \geq 0 \\
-1, & \mathbf{w}^\top \mathbf{x} < 0
\end{cases}
\]

or, more concisely:

\[
\hat{y} = \text{sign}(\mathbf{w}^\top \mathbf{x})
\]  

(5.1)

where, again, the last element of \( \mathbf{x} \) is set to 1 and the last element of \( \mathbf{w} \) is \(-t\).

The perceptron learns the weights \( \mathbf{w} \) using the so-called perceptron learning algorithm, or PLA, which is a supervised, gradient-based learning method (Bishop, 2006). To derive PLA, we start by recognizing that the perceptron is trying to place all points that belong to class +1 on one side of the hyperplane defined by \( \mathbf{w} \) and all points that belong to class \(-1\) on the other. This can be written as \( \mathbf{w}^\top \mathbf{x}_n > 0 \) for points \( \mathbf{x}_n \) that have true label \( y_n = +1 \) and as \( \mathbf{w}^\top \mathbf{x}_n < 0 \) for points that have true label \( y_n = -1 \). More concisely, we can write this as \( \mathbf{w}^\top \mathbf{x}_ny_n > 0 \). According to the so-called perceptron criterion, the perceptron receives no penalty for correctly classified points - regardless of how far they are from the boundary - and incurs a cost

\[
E(\mathbf{w}; \mathbf{x}_n, y_n) = -\mathbf{w}^\top \mathbf{x}_ny_n
\]  

(5.2)

proportional to how badly misclassified points fall on the wrong side of the boundary. This cost is what PLA aims to minimize. Mathematically, this can be written as

\[
E(\mathbf{w}) = \sum_{n=1}^{N} E(\mathbf{w}; \mathbf{x}_n, y_n)
\]

\[
\mathbf{w}^* = \arg \min_{\mathbf{w}} E(\mathbf{w})
\]

As long as the data are linearly separable, solutions \( \mathbf{w}^* \) will exist and can be achieved by stochastic gradient descent. This is involves sampling a misclassified
point, computing the gradient of the error with respect to $w$ for that point, and moving $w$ away from that direction. Mathematically, we can write this as:

$$w^{(i)} = w^{(i-1)} - \eta \nabla_w E(w; x_n, y_n)$$

$$= w^{(i-1)} + \eta x_n y_n$$

(5.3)

where the index $i$ denotes the iteration of PLA. $\eta$ is called the learning rate and controls how dramatically the weights are adjusted. If $\eta$ is too large, the weights will change too rapidly and the algorithm may not converge. On the other hand, if $\eta$ is too small, learning can be slow. Thus, in many practical applications, it is important to optimize the learning rate to balance speed with convergence rate.

Graphically, one can imagine PLA as shifting the weights of the perceptron such that the inner product between the weight and data points in the training set changes so that the $w^T x$ is “more correct” (Abu-Mostafa et al., 2012). For example, we can represent $w$ and $x$ as vectors in $\mathbb{R}^2$. If $\hat{y} = 1$ while $y = -1$, this means that $w$ and $x$ form an acute angle when they should be obtuse. An acute angle would result in a positive inner product since $w$ and $x$ would be pointing in the same general direction while an obtuse angle would result in a negative inner product since $w$ and $x$ would be pointing in opposite (in the worst case, antiparallel) directions. By subtracting $x$ from $w$, we make the angle more obtuse. This and the converse scenario are illustrated in figure 5.1.

In this project, we implement PLA for a perceptron completely in electronic hardware, giving our device the ability to function as a standalone computer that processes information as it is received. Together with the neuromagnetic reactor, the system comprises a hybrid analog-digital system, loosely mimicking the brain. There are important tradeoffs to consider between software and hardware implementations of a perceptron. Software can be completed in a handful of lines of code, as shown below, and is typically easier to modify than hardware. For instance, there are many
(a) Modification of weights for misclassified point in class $-1$. Originally, the angle between the weight and the data point is acute, resulting in a positive inner product and a prediction of $+1$. By modifying the weights according to equation 5.3, PLA yields a new weight which forms an obtuse angle with $x$. Thus, $x$ is now correctly classified.

(b) Modification of weights for misclassified point in class $+1$. Originally, the angle between the weight and the data point is obtuse, resulting in a negative inner product and a prediction of $-1$. By modifying the weights according to equation 5.3, PLA yields a new weight which forms a less obtuse angle with $x$. Although $x$ is not correctly classified after one PLA iteration, repeated applications of PLA result in a correct classification for $x$.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure5.1}
\caption{This illustration shows how the perceptron modifies its weights so that misclassified points are “more correctly” classified. As long as the data is separable, with enough training iterations, PLA will yield a set of weights that perfectly separates the data belonging to the two classes.}
\end{figure}

Different ways that operations such as summation, multiplication, and thresholding can be performed (e.g. via operational amplifiers, transistors, digital logic, etc.).

When choosing among these, one must consider tradeoffs between ease of implementation and performance / stability. Furthermore, when choosing components, one must be aware of device specifications to ensure that all inputs and outputs are within a tolerable range. Soldering components into a circuit can also be time-consuming and difficult to debug. If done well, however, a hardware implementation can result in substantial improvements in speed since operations can be performed in parallel.

% Vivek Subramanian
% Perceptron learning algorithm (PLA)
%% Inputs: x is N x D, y is N x 1, eta is a scalar
%% Outputs: w is D x 1

function w = Perceptron(x, y, eta)

w = randn(size(x, 2), 1);
thresh = randn;

yhat = x*w > -thresh;
misclass = yhat ~= y;
cond = sum(misclass) > 0;

while cond
    fmisclass = find(misclass);
    updateidx = fmisclass(randi(length(fmisclass), 1));
    sign = (-1*(y(updateidx) == 0) + (y(updateidx) == 1));
    w = w + eta*sign*x(updateidx, :);
    thresh = thresh + eta*sign;
    yhat = x*w > -thresh;
    misclass = yhat ~= y;
    cond = sum(misclass) > 0;
end

../Code/Perceptron.m

5.2 Hardware implementation of perceptron

Here, we describe a hardware implementation of the perceptron we developed which serves as the digital component of our hybrid computing device. The hardware perceptron contains four main components, which are described in detail in the following subsections. These are the timing circuit, the data input and prediction circuit, the phase identifier, and the learning circuit. A timing circuit controls when inputs will be received and processed; a data input and prediction circuit reads in new data and determines to which class it belongs; a phase identifier determines which state of processing to perform within a clock cycle; and the learning circuit carries out the analog arithmetic to allow the perceptron to learn. A dual +5 V and −5 V power supply powers the circuit. A full diagram of the hardware perceptron is given in figure 5.2.
Our perceptron is a working prototype, and we have simplified our design by limiting the model to process a single scalar input and to have a single parameter. This parameter \( t = w_2 \) is a threshold which causes the device to output \text{high} if the input is above the threshold or \text{low} otherwise. With this constraint, we can expand \( w \) as \( w = [1 \quad -t]^T \) and \( x_n \) as \( x_n = [x_n \quad 1]^T \). The prediction of the perceptron and the cost function are still given by equations 5.1 and 5.2, but now, because \( w_1 \) is fixed to 1, we find that:

\[
\begin{align*}
  w_1^{(i)} &= w_1^{(i-1)} - \eta \nabla_{w_1} E \\
  &= w_1^{(i-1)} - \eta \nabla_{w_1} [(1)(x_n)(y_n)] \\
  &= w_1^{(i-1)}
\end{align*}
\]

Thus, we only need to update the parameter for \( w_2 \) which corresponds to the negative threshold \(-t\):

\[
\begin{align*}
  w_2^{(i)} &= w_2^{(i-1)} - \eta \nabla_{w_2} E \\
  &= w_2^{(i-1)} - \eta \nabla_t [(-t)(1)(y_n)] \\
  &= w_2^{(i-1)} + \eta y_n
\end{align*}
\]  

(5.4)

Note that the update for \( w_2 \) does not depend on the value of \( x_n \) since the second element of \( x_n \) is 1. At the end of this chapter, we briefly outline a method by which we could extend our implementation to allow for more weights.

5.2.1 Timing circuit

A circuit diagram of the timing circuit is shown in figure 5.3. The part of the circuit outlined in red is a Schmitt trigger, which is an operational amplifier (op amp) wired as a comparator with positive feedback. The output of the Schmitt trigger switches to the upper rail of the op amp (\text{high}, here +5 V) when the inverting input falls below a lower threshold. Similarly, the output is the lower rail (\text{low}, here 0 V) when
Figure 5.2: Simplified perceptron implemented in hardware. The perceptron consists of four main circuits to control timing (maroon), data input and prediction (yellow), phase detection (orange), and learning (purple) and is powered by a dual +5 V and −5 V power supply (white). The bright red and lime green boxes are specific components of the timing and phase detector circuits that will be discussed in further detail.
Figure 5.3: Timing circuit consisting of a Schmitt trigger (outlined in red) which charges and discharges a capacitor. The time constant of the capacitor dictates the speed of computations of the perceptron.

the inverting input rises above an upper threshold. The thresholds are set by the reference voltage (here also +5 V) and a set of three resistors sharing a node at the non-inverting input to the op-amp. Refer to figure 5.4 for a detailed description of how these resistors are chosen. The op amp used to build this Schmitt trigger is a Texas Instruments MC1458P; it is suitable for low power applications (able to supply current of up to 25 mA), compatible with the ±5 V power supply, and comes in an eight 8-pin dual inline package that allows us to save space on our final board design (Texas Instruments, 1971).

The output of the Schmitt trigger, labeled \( \text{chg} \), in figure 5.3 is used to switch a pair of photorelays \( S_4 \) and \( S_5 \) which control whether the capacitor \( C_{\text{timing}} \) charges or discharges. In the schematic, \( S_4 \) is of type SW and is normally open, meaning that when the control voltage \( \text{chg} \) is low, the relay acts as an open circuit (and
Figure 5.4: Schmitt trigger with asymmetric thresholds. This is essentially a comparator with positive feedback. Assume $V_{\text{ref}} = V_{\text{DD}}$. Thresholds are calculated by noting that when the output goes to $V_{\text{DD}}$ (noninverting input voltage > inverting input voltage), $R_1$ and $R_2$ are in parallel. Thus, the lower threshold is given by $\frac{R_3}{R_3 + R_1} V_{\text{ref}}$, where $R_1 || R_j = \frac{R_i R_j}{R_i + R_j}$. Similarly, when the output goes to 0V (inverting input voltage > noninverting input voltage), $R_2$ and $R_3$ are in parallel. Thus, the upper threshold is given by $\frac{R_2 || R_3}{R_2 + R_3} V_{\text{ref}}$. One can then choose $R_1$, $R_2$, and $R_3$ based on the desired thresholds.

vice versa). On the other hand, $S_5$ is of type SW2 and is normally closed, meaning that when $\text{chg}$ is low, the relay is a short circuit (and vice versa). The photorelays used in this design are the Toshiba TLP222AF (normally open) and the Panasonic AQY412EH (normally closed) (Toshiba, 2017; Panasonic, 2017). Each of these is current-controlled: the TLP222AF switches closed for currents between 5 and 25 mA while the AQY412EH switches open for currents between 5 and 30 mA. Moreover, the TLP222AF has a typical dropout voltage of 1.15 V while the AQY412EH has a typical dropout voltage of 1.25 V. Since our Schmitt trigger has a supply (rail) voltage of 5 V, we can achieve a forward current in the desired range using resistors $R_{48}$ and $R_{49}$ that have values between 154 and 770 Ω for the TLP222AF and between
150 and 750 Ω for the AQY412EH. Thus, we chose a 270 Ω resistor which falls within this range.

The resistors $R_{45}$, $R_{47}$, and $R_{46}$ corresponding to $R_1$, $R_2$, and $R_3$ in figure 5.4 were chosen to be 1 kΩ, 250Ω, and 20 kΩ, respectively. This choice results in lower and upper thresholds of 1 V and 4.95 V, respectively. Initially, $V_{\text{timing}}$ is 0 V. This causes the op amp $U_4$ to output high, thus closing $S_4$ and opening $S_5$. $C_{\text{timing}}$ begins charging, but when its voltage reaches 4.95 V, the Schmitt trigger output switches to the lower rail. This causes $S_4$ to open and $S_5$ to close. $C_{\text{timing}}$ discharges across $R_2$, and when $V_{\text{timing}}$ reaches 1 V, the Schmitt trigger output switches back to the upper rail. This causes $S_4$ to close and $S_2$ to open, allowing $C_{\text{timing}}$ to charge once again.

The time constant of the RC circuits consisting of (1) $C_{\text{timing}}$ and $R_6$ and (2) $C_{\text{timing}}$ and $R_7$ dictate how quickly the device processes a single data point. In this case, both $R_6$ and $R_7$ are the same value of 1 kΩ, and $C_{\text{timing}}$ has a value of 4.7 µF. This means that the time constant of both circuits is $\tau = RC = (1 \text{ kΩ})(4.7\mu F) = 4.7$ ms. Thus, 4.7 ms are required for $C_{\text{timing}}$ to charge up from 0 V to $(1-1/e)V^+ = 0.63V^+$ V and to discharge from $V^+$ to $V^+/e$ V, where $e \approx 2.718$ is a constant. (Five time constants are typically required for a capacitor to completely charge or discharge in an RC circuit.) Whenever $V_{\text{timing}}$ reaches 2 V during its rising phase, a new data point is fed to the device. Thus, the timing circuit provides a clock signal, and different operations are performed during the rising and falling phases of the clock cycle. These operations take place during the four phases of the phase identifier, discussed in section 5.2.3. Without the timing circuit, learning would have to be implemented asynchronously; this would require sending a “done” signal to the remainder of the circuit when the CPU is finished learning from a data point. Even with slight delays in the arrival or transmission of a signal - for instance, if subsequent
data points arrive before the current data point has finished being processed - learning could be thrown off. To avoid these difficulties, and in favor of rapidly creating a working prototype, we chose to implement a synchronously timed perceptron.

The smaller the time constant, the more data points that can be processed in a given amount of time. However, making the time constant too small can result in errors because the photorelays may not have enough time to switch on or off. This will be addressed again when the results are presented, but we briefly touch on the issue here. Photorelays have a characteristic known as turn on time, which represents the time required for the light from the LED to be detected by the photodetector and close/open the circuit. A similar characteristic called turn off time also exists to perform the opposite function. From the datasheet for the TLP222AF, we find that its turn on time has a maximum value of 2 ms. For the AQY412EH, we find that it has a maximum turn on time of 10 ms; its typical value, however, is much less - only 3 ms. The turn off times for each of the devices is much less (0.5 ms and 1 ms maximum for the TLP222AF and the AQY412EH, respectively). As will be discussed below, we designed an LTspice simulation prior to building out the physical circuit. Our LTspice simulation incorporates the relays but does not factor in the turn on and turn off times. While the simulation behaved as expected since the circuit had an idealized turn on and turn off time of 0 ms, the actual circuit had a non-negligible turn on and turn off time. Although the physical circuit behaved as expected for inputs we tested and the long term behavior of the device was satisfactory (i.e. the average weight learned as a function of time was acceptable), some updates were missed. We believe that this was due to operating the photorelays outside the optimal range. Future iterations of the device will incorporate photorelays with shorter turn on times.
5.2.2 Data input and prediction

New data points \((x_n, y_n) : n \in 1, \ldots, N\), where \(N\) is the total number of data points, are fed into the circuit when \(V_{\text{timing}}\) is rising and reaches 2 V. (Hereafter, depending on context, we will refer to \(x_n\) and \(y_n\) as both the actual data point and the voltage by which they are encoded.) These data points are encoded as a single voltage whose value is maintained until \(V_{\text{timing}}\) reaches 2 V during its next rising cycle, at which time a new data point is fed in and the voltage is changed. As soon as the data point is inputted, the perceptron makes a prediction and determines whether an update needs to be performed. This can be understood by inspecting the subcircuit depicted in figure 5.5.

If the prediction of the perceptron matches the label of the data point, no changes need to be made to the weight of the perceptron. On the other hand, if the prediction and the label conflict, the weight needs to be updated to reduce the chance of making subsequent errors. The update that is performed is known as the perceptron learning rule and is given by equation 5.3. In our hardware, the process of updating the weight takes place during four phases, and the timing of these phases is dictated by the phase identifier subcircuit. This is essentially a group of window comparators that acts in parallel on \(V_{\text{timing}}\): when the voltage \(V_{\text{timing}}\) across \(C_{\text{timing}}\) falls within a certain range, the computation associated with that window is executed. In the next subsection, we outline the four phases and explain how the window comparator is applied to trigger the start and stop times of each of these phases.

5.2.3 Phase identifier

As \(C_{\text{timing}}\) charges and discharges, its voltage \(V_{\text{timing}}\) is fed into another subcircuit - the phase identifier. The phase identifier consists of four nearly identical subcircuits which differ only in resistor values. One of these subcircuits is depicted in figure 5.6. To the very left of the figure is a part of the phase identifier that is not repeated.
Figure 5.5: Subcircuit that is responsible for performing predictions. When a new data point \( x_n \) arrives, its value is inverted by \( U_{16} \) and fed into the inverting input of the comparator \( U_1 \). The photorelay S8 is normally closed so the voltage \( W \) across \( C_{\text{weight}} \) is fed into the noninverting input of the comparator \( U_1 \) and compared to \( x_n \). If \( x_n > -V(C_{\text{weight}}) \), the comparator outputs high, and vice versa. Recall that by definition of the perceptron, the rule that dictates the true label of the data is given by equation 5.1, where the last element of \( w \) is \(-t\), the negative threshold. But the PLA rule updates \( w \), and in our case, \( w \) reduces to \( w = -t \) since we have only one parameter. Hence, when we learn \( w \), we are actually learning the negative of the threshold. Thus, to get the threshold which we typically interpret (and compare with \( x \)), we simply flip the sign of the \( w \) that we learn. Finally, the XOR logical gate \( A_1 \) determines whether the prediction - i.e. the output of \( U_1 \) - matches the label \( y_n \) of \( x_n \). If the prediction and label match, \( A_1 \) outputs low and vice versa.

in each of the four components. In this branch, \( V_{\text{timing}} \) is first buffered and then passed through a high pass filter comprised of \( C_4 = 4.7\mu F \) and \( R_8 = 500\Omega \). The cutoff frequency of this high pass filter is therefore:

\[
 f_{\text{cutoff}} = \frac{1}{2\pi RC} = \frac{1}{2\pi(500\Omega)(4.7\mu F)} = 67.7 \text{ Hz} \tag{5.5}
\]

The high pass filter also acts as a differentiator, outputting a value proportional to the rate of change of the input voltage. If the cutoff frequency is too high (\( R \) and/or \( C \) too small), the amplitude of the response will be too weak due to excessive filtering.
Figure 5.6: The phase identifier circuitry consists of four window comparators like the one shown in this figure. The comparators determine whether the $V_{\text{timing}}$ is within a given range, and if so, provide an output signal which is utilized by the computing circuit to perform a phase of the update. The part of the figure in the green box consists of a high pass filter / differentiator followed by a comparator. This ensures that the phase identifier is only active in modifying the weight $W$ when $V_{\text{timing}}$ is discharging.

If the cutoff frequency is too low ($R$ and/or $C$ too large), not enough filtering will occur and the entire input signal will pass. We performed a simulation in LTspice to determine values of $R$ and $C$ that would yield the correct slope. The result of the simulation for the choices given above are shown in figure 5.7.

The output of the high pass filter is fed into a comparator $U_5$ which outputs high if the input is less than 0 V (ground). This tells the remainder of the circuit that $C_{\text{timer}}$ is discharging. The four phases of weight update take place during this time interval. These phases are triggered by a set of window comparators. One of these window comparators is shown by the part of 5.6 not outlined in green; this block is repeated four times, and together with the green block, the four are shown in orange.
(a) RC high pass filter with cutoff frequency of 67.7 Hz.

(b) Response of circuit to exponential saturation and decay.

**Figure 5.7:** The input voltage (shown in blue in figure 5.7b) used to test this RC high pass filter mimics that of a capacitor undergoing a single charge and discharge cycle. The time constant of 4.7 ms is the same as the time constant of $V_{\text{timing}}$. Since the output (shown in green in figure 5.7b) yields a voltage that is proportional to the rate of change of $V_{\text{timing}}$ and is large enough to provide a useful signal, we proceeded with $C_4 = 4.7 \mu F$ and $R_8 = 500 \Omega$.

Each window comparator consists of three resistors which control the upper and lower thresholds of the window. They act as voltage dividers, splitting the reference voltage $V^+ = 5 \text{ V}$ into two parts. Consider the first window comparator consisting of the two regular comparators $U_6$ and $U_7$ and the three resistors $R_{12}$, $R_{10}$, and $R_{11}$. The resistors divide $V^+$ into two voltages $V_a = \frac{R_{10} + R_{11}}{R_{12} + R_{10} + R_{11}} V^+$ and $V_b = \frac{R_{11}}{R_{12} + R_{10} + R_{11}} V^+$. $U_6$ receives as its inputs $V_{\text{timing}}$ and $V_a$ and outputs high if $V_{\text{timing}} < V_a$; $U_7$ receives
as its inputs $V_{\text{timing}}$ and $V_b$ and outputs high if $V_b < V_{\text{timing}}$. The AND logic gate $A_2$ can be considered a high input impedance device, meaning that very little current flows into it. Labeling the outputs of $U_6$ and $U_7$ as $V_6$ and $V_7$, respectively, and the node between the two $1\,\text{k}\Omega$ resistors as $V_c$, by Kirchoff’s rules, we find:

$$\frac{V_6 - V_c}{1\,\text{k}\Omega} = \frac{V_c - V_7}{1\,\text{k}\Omega}$$

$$V_6 - V_c = V_c - V_7$$

$$V_c = \frac{V_6 + V_7}{2}$$

Since $V_6$ and $V_7$ are opposite rails when $V_{\text{timing}}$ falls outside the range of $V_a$ and $V_b$, their sum goes to 0. On the other hand, $V_6$ and $V_7$ are both 5 V when $V_{\text{timing}}$ is between them, so the output is 5 V. Thus, the output of the combined system is high only when $V_b < V_{\text{timing}} < V_a$; otherwise the output is zero. The remaining window comparators function analogously.

The specific choices of resistors for each of the window comparators results in the following four windows: 4.25 V to 3.75 V, 3.5 V to 3.00 V, 2.75 V to 2.25 V, and 2.00 V to 1.5 V. Recall that when $V_{\text{timing}}$ falls to 1 V, the timing circuit switches back to charging $V_{\text{timing}}$ until it reaches 4.95 V, at which point $V_{\text{timing}}$ begins to discharge again. Thus, the four windows fall in between the two limits 1 V and 4.95 V of the Schmitt trigger in the timing circuit. Also recall that, in the prediction circuit, when there is a mismatch between the predicted label and truth, the XOR gate $A_1$ outputs high. The output node, labeled inc, which takes on a value of $+5V$ or $0V$, feeds into the four AND gates, one for each phase in the phase identifier. When inc is high and one of the window criteria is met, the output of the corresponding AND gate also goes high. This triggers one of the four phases - titled reset1, compute, reset2, and update - which are identified by the node labels of the outputs of the
AND gates. Note that only one phase is engaged at a time because the windows of the window comparators are non-overlapping.

5.2.4 Learning circuitry

We have explained how the computations are timed using a combination of the timing circuit and phase identifier, and now we describe the mechanics of each of the four phases and how they accomplish PLA. The learning circuitry is depicted in figure 5.8.

During each of the four phases, photorelays are activated which close parts of the circuit to perform a computation. Notice that all of the photorelays in this circuit except $S_8$ are the normally open type, corresponding to the TLP222AF; $S_8$ is normally closed (AQY412EH). When $4.25\, V > V_{\text{timing}} > 3.75\, V$, the reset1 phase is active, and the voltage labeled reset1 is high. $S_{10}$ and $S_{11}$ close, causing $C_{\text{temp}}$ to discharge across $R_{26} = 0.1\, \Omega$. The time constant of the discharge is given by:

$$\tau = R_{26}C_{\text{temp}} = (0.1\, \Omega)(0.5\mu F) = 50\, \text{ns} \quad (5.6)$$

Thus, relative to the time for which $S_{11}$ and $S_{10}$ are closed (on the order of ms), the discharge is instantaneous. The phase is called reset1 because the temporary voltage-storing capacitor $C_{\text{temp}}$ is reset during this phase. At the end of the reset1 phase, $S_{11}$ and $S_{10}$ go back to being closed.

When $3.50\, V > V_{\text{timing}} > 3.00\, V$, the compute phase is active, and the voltage labeled compute is high. $S_3$, $S_6$, and $S_2$ close. This allows the voltage on the weight-storing capacitor $C_{\text{weight}}$ to appear at the buffer amplifier $U_{15}$. The buffer converts the high impedance source into a low impedance source and feeds the voltage into the inverting summing amplifier $U_2$. On the left hand side of the circuit, we see that a voltage divider is being applied to the input $y_n$. This voltage divider implements the learning rate for PLA. We chose a learning rate of 0.1, so to achieve this, we use a $90.9\, k\Omega$ resistor in series with a $10\, k\Omega$ resistor. Thus, the input to the buffer
The learning circuitry consists of a network of photorelays controlling the charging and discharging of two capacitors. One capacitor $C_{\text{weight}}$ stores the weight parameter of the perceptron. The other $C_{\text{temp}}$ stores a voltage temporarily and is used to update the voltage held by $C_{\text{weight}}$. The learning circuit is the part of the perceptron that implements the PLA given by equation 5.4. The learning rate $\eta$ is controlled by two resistors $R_4$ and $R_3$ which act as a voltage divider, scaling the value of the label $y_n$. $U_2$ acts as an inverting summing amplifier with a gain of 2; thus, the scaled label voltage is added to the existing value of $V_{w}$, inverted, and then doubled. This doubled voltage is eventually applied across $C_{\text{temp}}$ to charge the capacitor, and later, since the capacitances of $C_{\text{temp}}$ and $C_{\text{weight}}$ are equal, $C_{\text{temp}}$ equilibrates with $C_{\text{weight}}$ with both reaching half of $V(C_{\text{temp}})$ in magnitude but opposite in sign. More details are provided in the text.
amplifier $U_3$ is given by $\frac{10 \times 10^3 \Omega}{10 \times 10^9 + 90.9 \times 10^7} y_n$ and represents the mathematical operation $\eta y_n$. The buffered signal is then sent into the inverting summing amplifier $U_2$ as its second parallel input. The operation of the summing amplifier is described in detail in figure 5.9; essentially, the voltages appearing in parallel at its inverting input are summed and a gain given by the negative ratio of the resistance of the feedback resistor to that of one of the identical input resistors is applied to the sum. (If the parallel inputs are fed into the noninverting input, there is no negation.)

The output, denoted by $V_{\text{sum}}$ in the schematic, is then applied across the series combination of $C_{\text{temp}}$ and $R_{20}$. Recall that $C_{\text{temp}}$ was just discharged completely during the reset1 phase. Thus, charging of $C_{\text{temp}}$ to the voltage outputted by $U_2$ begins from 0 V and takes place almost instantly since the time constant of this circuit is identical to that of the discharging circuit in the reset1 phase. The ratio of $R_5 = 20 \, \text{k}\Omega$ to either $R_1 = 10 \, \text{k}\Omega$ or $R_2 = 10 \, \text{k}\Omega$ is 2, so as long as the output falls between the positive and negative supply voltages, it will be amplified by a factor of 2. This will play a role in updating $C_{\text{weight}}$ as explained below. Finally, at the end of the compute phase, the switches $S_3$, $S_6$, and $S_2$ switch back to open.

Note that it is essential to store voltage on a temporary resistor rather than directly modifying the voltage on $C_{\text{weight}}$. This is because $C_{\text{weight}}$ provides one of the inputs to the inverting summing amplifier $U_2$; thus, if its value is changed while it is still wired to the amplifier, the summation will be affected. (Essentially, the voltage would saturate to one of the rails of the amplifier because of the feedback loop.)

Next, when $2.75 \, \text{V} > V_{\text{timing}} > 2.25 \, \text{V}$, the reset2 phase is active, and the voltage labeled reset2 is high. $S_8$, which is normally closed, opens, isolating the learning circuit from the rest of the perceptron. During this phase, the voltage across $C_{\text{weight}}$ will be reset, so to ensure the this does not cause conflicting predictions in
Because the op amp aims to keep the voltage at the two input terminals the same, the voltage at the inverting input is treated as a virtual ground. Furthermore, virtually no current enters the input terminals of an op amp because of their high input impedance. By applying Kirchoff’s current law, we find that the current entering $V_{\text{in}}^{-}$ equals the current leaving $V_{\text{in}}^{-}$. Thus, $\frac{V_1}{R} + \frac{V_2}{R} = -\frac{V_{\text{out}}}{R_g}$. Solving this for $V_{\text{out}}$, we find that $V_{\text{out}} = -\frac{R_g}{R} (V_1 + V_2)$.

More voltages can be applied at $V_{\text{in}}^{-}$ to sum even more signals, i.e. $V_{\text{out}} = -\frac{R_g}{R} (V_1 + V_2 + \cdots)$.

In addition, $S_9$ is closed; this allows $C_{\text{weight}}$ to discharge across $R_{17}$ (with the same time constant as before, making the discharge nearly instant). At the end of the reset2 phase, $S_8$ is again closed and $S_9$ is opened.

Finally, when $2.00 \text{ V} > V_{\text{timing}} > 1.50 \text{ V}$, the update phase is active, and the voltage labeled update is high. $S_7$ and $S_2$ are closed, placing the two capacitors “almost in parallel” with each other. Thus, $C_{\text{temp}}$ as as a source of while $C_{\text{weight}}$ is a sink for charge. When the two are connected, $C_{\text{temp}}$ discharges while $C_{\text{weight}}$ charges. The 0.1 $\Omega$ resistor is included to prevent the capacitors from charging / discharging too quickly and causing a spark, which is more likely if they were directly connected. It does not, however, affect the end voltage across each of the capacitors. Currently, $C_{\text{temp}}$ has been charged up to negative twice the voltage that $C_{\text{weight}}$ needs to reach. The negation is necessary because the source that charges $C_{\text{temp}}$ is (schematically) below $C_{\text{temp}}$ and ground is above $C_{\text{temp}}$. On the other
hand, $C_{\text{temp}}$, the source that is charging $C_{\text{weight}}$ is above $C_{\text{weight}}$, and ground is below. When these two capacitors are connected across each other, their polarities are reversed. Thus, by charging $C_{\text{temp}}$ with an inverting summing amplifier, we accommodate for this reversal. Since both capacitors have identical capacitances of $0.5\mu F$, the voltage reached by each capacitor after discharging / charging will be half the voltage of $C_{\text{temp}}$. This is why the gain factor of 2 in $U_2$ is necessary. At the end of the update phase, $S_7$ and $S_2$ are opened.

Note that after the four phases have completed, the voltage on $C_{\text{temp}}$ is still half the magnitude of the voltage on $C_{\text{weight}}$ and the opposite sign. Before another update can be performed, $C_{\text{temp}}$ must be reset so that output of $U_2$ charges it to the desired voltage, without any effect from the previous update cycle. As described above, this is accomplished by the reset1 phase; we mention it here for clarity. Figure 5.10 illustrates the update cycle schematically.

5.3 LTspice simulations

In this section we describe the results of circuit simulations performed in LTspice to verify the design would function as expected. We follow the same logical flow as above, first detailing the results of the timing circuit; then data input and prediction; followed by the phase identifier; and finally, the learning circuit.

Figure 5.11 shows the voltage $V_{\text{timing}}$ which dictates the timing of the perceptron and its control voltage $\text{chg}$. As expected, $\text{chg}$ is initially high ($+5$ V) since the $V_{\text{timing}} = 0 < 4.95V$. $V_{\text{timing}}$ then charges from 0 V to 4.93 V in 20.39 ms. This is slightly under 5 time constants, the time needed to charge a capacitor from 0 V to 99.3% of the supply voltage: $\tau = RC = (4.7\mu F)(1 \text{ k}\Omega) = 4.7 \text{ ms} \rightarrow 5\tau = 23.5\text{ms}$. Moreover, when $V_{\text{timing}}$ reaches 4.93 V, the Schmitt trigger output changes from high to low. This causes $V_{\text{timing}}$ to discharge. Resistors were chosen so that the
Figure 5.10: The phase identifier consists of four phases: \texttt{reset1}, \texttt{compute}, \texttt{reset2}, and \texttt{update}. Each of the phases results in a change in the voltage a temporary weight storage capacitor \(C_{\text{temp}}\) and/or the permanent weight storage capacitor \(C_{\text{weight}}\).
lower threshold of the Schmitt trigger was 1 V. We see that at 27.60 ms, when $V_{\text{timing}}$ reaches 1.08 V, $\text{chg}$ switches back from low to high, and $V_{\text{timing}}$ begins charging again. This cycle of discharging from 4.93 V to charging at 1.08 V repeats until the end of the simulation.

Next, we show the data input and prediction results. For data input to the simulation, the simulated $V_{\text{timing}}$ was first imported into MATLAB and to determine how many data points would have to generated. This was done by counting the number of rising phases of $V_{\text{timing}}$ during the total simulation duration of 2.5 s. 96 rising phases were identified, so a time series was created such that at the trigger point of 2 V during the rising phase of $V_{\text{timing}}$, a new data point would be introduced. Its value would be constant until the next trigger point. Both $x$ and $y$ coordinates were created at once, and a threshold separating data points belonging to class $-1$ from those belonging to class $+1$ was arbitrarily decided. Several different choices of threshold were tested and all attempts successfully allowed the perceptron to learn.
We report on a simulated dataset in which all data points belonging to class $-1$ fell between 0.74 and 0.86; all data points belonging to class $+1$ fell between 0.93 and 1.05. Thus, a valid threshold falls between 0.86 and 0.93. The time series was written in a format that could be imported into LTspice. MATLAB code for this process is provided below.

```matlab
clear; close all

% Load time series exported from LTspice
M1 = importdata('Schmitt_trigger_oscillator.txt');

time_lt = M1.data(:, 1);
Vcap_lt = M1.data(:, 2);

% Find the times for which capacitor is charging and voltage is 2.0 V
dVcap_lt = [0 diff(Vcap_lt)];
new_input_lt = [0 diff((Vcap_lt > 2 & Vcap_lt < 2.5) & (dVcap_lt > 0))] > 0;

f_new_input_lt = find(new_input_lt);
first_input_time_lt = time_lt(f_new_input_lt(1));
interval_lt = mean(diff(time_lt(f_new_input_lt)));

figure; plot(time_lt, Vcap_lt, 'k')
hold on; plot(time_lt, new_input_lt, 'r')
xlabel('Time (s)')
ylabel('V_{cap}')
legend('V_{cap}', 'new input')

% Generate random samples that will be fed in at those times
rng(1)
mu1 = 0.8;
mu2 = 1;
if mu1 >= mu2
    error('mu1 must be < mu2.')
end
n = ceil(length(f_new_input_lt)/2) * ones(1, 2);
x = [mu1 + 0.03*randn(n(1), 1); mu2 + 0.03*randn(n(2), 1)];
y = [-1*ones(n(1), 1); ones(n(2), 1)];

figure; scatter(x(y == -1, 1), zeros(length(x(y == -1)), 1), 'ko');
hold on;
delay = x(y == 1, 1), zeros(length(x(y == 1)), 1), 'ro';

% shuffle the data points before creating the time series
idx = randperm(length(x), length(x));
```

182
Figure 5.12: Simulation results for data input circuitry showing \( x \) and \( y \) time series. As expected, data change points occur when \( V_{\text{timing}} \) is on the rising phase and reaches 2.01 V.

\begin{verbatim}
x = x(idx);
y = y(idx);

%% Generate PWL text file

%% Uses the start time and interval calculated above.
fidX = fopen('myInputX5V.txt', 'w');
fidY = fopen('myInputY5V.txt', 'w');
fprintf(fidX, '0\t0\n%0.9f\t0\n', first_input_time_lt);
fprintf(fidY, '0\t0\n%0.9f\t0\n', first_input_time_lt);
for i = 1:length(f_new_input_lt)
    fprintf(fidX, '+.1u\t%0.9f\n+%0.9f\t%0.9f\n', x(i), interval_LT, x(i));
    fprintf(fidY, '+.1u\t%0.1f\n+%0.9f\t%0.1f\n', y(i), interval_LT, y(i));
end

../Code/Data_generator.m
\end{verbatim}

Figure 5.13 shows the value of \( x \) and \( y \) for the same time period in the simulation as was shown in 5.11. We see that, as expected, the input changes when \( V_{\text{timing}} \) is on its rising phase and reaches 2.01 V; it remains constant until the next trigger point. Furthermore, \( y \) assumes values of either -1 or +1 and \( x \) assumes continuous values. When \( x > 0.93 \), \( y = +1 \); when \( x < 0.86 \), \( y = -1 \).
We next verify the prediction circuitry. Predictions are made by comparing $-x$ to the threshold $V_w$. For demonstration, we show in figure 5.13 one cycle in which there was a mismatch between prediction and truth resulting in an update for $V_w$. There are six plot panes in the figure. The top plot pane shows $V_{\text{timing}}$. The value of $V_{\text{timing}}$ whose function has already been verified. The second pane title $V(\text{out2})$ shows $-x$, which is the output of the inverting buffer $U_{16}$. When $V_{\text{timing}}$ reaches approximately 2 V on its charge cycle, we see that a new data point has been inputted, as $-x$ has changed. The third pane shows the voltage to which $-x$ will be compared - the threshold voltage $V_w$. We see that $-x \approx -780$ mV which is less than $V_w \approx -594$ mV. This means that the output of the comparator $U_1$ should be high; this expectation matches the simulation, which shows the output of $U_1$ as $V(n003)$ in plot pane 4. Next, we determine whether the true label matches the prediction. From plot pane 5, which shows $y$, we find that the label for this data point is $-1$. Thus, the prediction is high while the truth is low. This mismatch means that the prediction was incorrect; hence the node inc which is the output of the XOR gate $A_1$ is high. We also see that at around 263 ms, the weight $V_w$ is adjusted because of the mismatch.

We now verify that the phase identifier is working as expected. Figure 5.14 contains two panes. The top pane shows signals related to the high pass filter / differentiator, which determines when $V_{\text{timing}}$ is in the discharging part of its cycle. The gray trace shows $V_{\text{timing}}$ and the blue trace titled $V(n035)$ shows the output of the differentiator. As we saw previously, the passive differentiator does a relatively good job capturing the slope of $V_{\text{timing}}$. The red trace titled $V(n048)$ in the same pane shows the output of the comparator $U_5$, which, as it should, outputs high when the slope is less than 0. However, by comparing $V(n048)$ to $V_{\text{timing}}$, we find that there is a part of $V(n048)$ that remains even after $V_{\text{timing}}$ has started to increase again. This is because the differentiator does not charge and discharge quickly enough.
Figure 5.13: Plots show voltages relevant to the prediction circuitry. We find that while $-x$ is below the threshold, making the prediction high, the actual label for the data point is low. An adjustment is made to $V_w$ per PLA.

(i.e. its time constant is a bit too long). We experimented with smaller time constants and found that while the slope was identified more precisely when a smaller time constant was used, the amplitude of the slope also decreased. Because we knew that the overhang would not affect the performance of the circuit, we chose to keep the longer time constant with the greater amplitude since, in practice, having low signal can get washed out by noise and the circuit may not function correctly even if, in theory, it did.

The second pane of figure 5.14 shows the phase identifier with $V_{\text{timing}}$ shown again for reference. Recall that this was a data point for which there was a mismatch between prediction and truth; thus, inc is high. We find that, as expected, when $V_{\text{timing}}$ is falling and its value is between 4.25 V and 3.75 V, reset1 goes high. The other three phases behave similarly (measurements not shown on figure). Notice that there is one anomalous window that occurs when $V_{\text{timing}}$ is rising. This is due to the overhang in the output of the comparator $U_5$. As we will see in the discussion of the learning circuitry, it does not have any impact on the result.
Figure 5.14: Simulation results for phase identifier. The sign of the slope of $V_{\text{timing}}$ calculated by $U_5$ and titled as $V(n048)$ in the upper pane is mostly correct except for a slight overhang when $V_{\text{timing}}$ starts to rise. This has no impact on performance. The four phases are correctly detected by the window comparators as shown by the appropriately timed high voltage outputs of reset1, compute, reset2, and update. There is one additional time during which update is detected due to the overhang in $V(n048)$, but again, this has no impact on performance.

Finally, we discuss the simulation of the learning circuitry. We divide this analysis into four separate figures to avoid clutter. First, in figure 5.15, we show that during reset1, the voltage on $C_{\text{temp}}$ is erased. Prior to the start of reset1, the voltage on $C_{\text{temp}}$ is -594 mV. During reset1, it drops to -1.24 nV (essentially 0 V). Note that here, we are subtracting the voltage of the node below $C_{\text{temp}}$ from the voltage above it to highlight that $V(C_{\text{temp}})$ and $V(C_{\text{weight}})$ initially start of at the same value at every update; both are measured by subtracting the voltage on the node connected to virtual ground from the node they share when $S_2$ is closed.

Second, in figure 5.16, we show that during the compute phase, the voltage on $C_{\text{temp}}$ charges up to the negative twice the sum of $\eta y_n$ and $V_w$. The top pane shows the start and end times of the compute phase. The middle pane shows the addition. We know from figure 5.14 that $y_n$ for this data point is $-1V$; thus, $V(\text{in1}),$
which is $\eta y_n$, correctly shows $-100 \text{ mV}$. This is the first input to the inverting summing amplifier $U_2$. The second input should be equal to $V_w$. This is titled $V(\text{in2})$ in the middle pane, and by inspection, we see that they are equal during the compute phase. Note that the only thing separating $V_w$ and $V(\text{in2})$ is the unity gain buffer. The value of $V_w$ is $-594 \text{ mV}$, so at the output of $U_2$, we should expect to see $V_{\text{sum}} = -2(-0.594 - 0.1) = 1.388 \text{V}$. The trace titled $V(\text{sum})$ shows the output of $U_2$ which does indeed match our expectations at a measured voltage of $1.38 \text{V}$. Finally, we should see this voltage appear across the $C_{\text{temp}}$. This is shown in the bottom pane, in which the $V(C_{\text{temp}})$ rises from $0 \text{ V}$ to $1.38 \text{V}$. Note that here, unlike in the previous plot, we are subtracting the voltage of the node above $C_{\text{temp}}$ from the voltage below it because we want to emphasize that the capacitor is being charged to the same voltage as the output of $U_2$.

Third, in figure 5.17, we show that during reset2, the voltage $V_w$ across $C_{\text{weight}}$ is erased. Prior to the start of reset2, the voltage on $C_{\text{temp}}$ is $-594 \text{ mV}$. During reset1, it drops to $-727 \text{ pV}$ (essentially $0 \text{ V}$). Note that before reset1 and reset2, $V(C_{\text{temp}})$ and $V(C_{\text{weight}})$ are equal.
Figure 5.16: compute phase of learning circuitry. The existing value of $V_w$ is summed together with $\eta y_n$ per PLA, and twice the negative of the result is stored in $V(C_{\text{temp}})$. Negation and amplification are necessary operations to perform PLA based on the design of this circuit (explained in more detail in the text).

Figure 5.17: reset2 phase of learning circuitry. $V(C_{\text{weight}})$ is reset, priming it for the update.
Finally, in figure 5.18, we show that during update, $V(C_{\text{temp}})$ equilibrates with $V(C_{\text{weight}})$. Just prior to the start of update, the voltage on $C_{\text{temp}}$ is -1.38 V (note that the sign has flipped because we are measuring the differential voltage as we did when analyzing reset1), and the voltage on $C_{\text{weight}}$ is 0 V. When update is initiated, $V(C_{\text{temp}})$ rises to -693 mV while $V(C_{\text{weight}})$ drops to -693 mV (measurement not shown). This is what we would expect since $V_w$ started at -594 mV, and the perceptron learning rule tells us that to update $w$, we should add $\eta y_n$ to its existing value; here $\eta = 0.1$ and $y_n = -1$, so $w^{t(i)} = w^{t(i-1)} + \eta y_n = -594 + (0.1)(1000) = -694$ mV.

Looking at the long term behavior of the circuit in figure 5.19, we see that $V_w$ converges to a reasonable value for the weight $w = -t$. After about 535 ms, we see that $V_w = -883$ mV. Recall that a valid threshold was deemed to be between 0.86 and 0.93 V. The value is close to the limit of 0.86 V because there is no penalty for margin between the threshold and the nearest data point. Thus, all values between 0.86 V and 0.93 V are equally good solutions for the perceptron. Also notice that as time progresses, $V_w$ slowly starts to increase; this is due to leakage currents that
dissipate energy stored in $C_{\text{weight}}$ across other high impedance loads such as the photorelays, op amps, and logic gates. The time constant of this decay is quite large (seconds, if not minutes), so it does not affect the performance of the perceptron for most practical applications.

5.4 Physical implementation: design choices

In this last section, we describe choices of components used for the physical implementation of the perceptron based on the circuitry discussed above.

5.4.1 Design choices and KiCad schematics

The assembled printed circuit board (PCB) of the final design is shown in figure 5.20; the associated schematics are given in figures 5.21 through 5.24. These schematics were created using the PCB design software KiCad. Note that the numbering of the components in the schematic of the physical design are different from the numbering done in the LTspice simulation.

We first bring to attention that new data points need to be fed into the hardware
whenever $V_{\text{timing}}$ reaches 2 V during its rising phase. To accomplish this, we first run the Schmitt trigger oscillator independently of the remaining circuit and record the $V_{\text{timing}}$. This time series can then be analyzed offline to determine when the change points occur, and accordingly, another two time series corresponding to $x$ and $y$ can be generated. Finally, we can input the pre-recorded time series as well as the generated $x$ and $y$ time series back into the perceptron circuitry to ensure the change points are properly synchronized with $V_{\text{timing}}$. We can see how this is accomplished by looking at figures 5.21 through 5.23. Note that in figures 5.22 and 5.23, $V_{\text{cap}}$, which corresponds to $V_{\text{timing}}$ in the LTspice simulation, is a local variable. Thus, the timing circuitry can operate independently of the remaining parts of the circuit. This means we can first turn on the device and record from $V_{\text{cap}}$ on pin 2 of $U_1$; then, after performing the aforementioned signal processing and data generation, we can feed the resultant signals as $V_{\text{cap}}$ to $U_2$ and $x$ and $y$ to $P_1$. This allows the inputs to be correctly synchronized with $V_{\text{timing}}$.

The prediction circuit also consists of op amps. Here, we chose to use a TCA0372,
Figure 5.21: PCB schematic for perceptron: top level overview.
Figure 5.22: PCB schematic for perceptron: timing circuit.
which is more suited to high-power applications as it is capable of driving 1 A of current (ON, 2013). While this is not necessary, we used it as a precautionary measure. The XOR gate employed to determine whether the prediction matched the output is a Texas Instruments CD74AC86 (Tex, 2003). With a supply of 5 V, this gate requires an input between 3.85 V and 5 V for the input to be considered high; similarly, voltages between 0 V and 1.65 V will be considered low. The output of the comparator $U_1$ will be either +5 or −5; since the low output of the comparator is outside the range of the low value accepted by the logic gate, we include a diode in the physical circuit to restrict the input to fall within the acceptable range. Note that this is not shown in the diagrams in figures 5.2 or 5.5; these schematics were made in LTspice which has logic gates whose inputs can be +5 V or −5 V.

Finally, the phase identifier and the learning circuitry also consists of op amps and AND logic gates. We apply both the MC1458P and TCA0372 op amps but the choice is arbitrary. For the AND gates, we employ the Texas Instruments SN74AC11 which accepts the same ranges of voltages for its high and low inputs. Because the output of the differentiator $U_5$ is either +5 V or −5 V, we again employ a diode in the physical implementation to restrict the inputs the and gates to fall within the desired range. Again, these design modifications are not shown in figures 5.2 or 5.6.

The simulated results for the perceptron conclude the results given in this chapter. We defer the experimental results for the final chapter in which the analog neuromagnetic reactor is combined with the perceptron to form a hybrid analog-digital computer. There, we will demonstrate that the hardware perceptron we have designed is capable of learning based on inputs that have been processed by the reactor. We also gently remind the reader that this perceptron has only a single weight but extension to multiple weights is straightforward. In general, the updates for each weight in a perceptron are linear and independent since by equation 5.3, we
Figure 5.23: PCB schematic for perceptron: data input and prediction circuit and phase identifier.
Figure 5.24: PCB schematic for perceptron: learning circuitry.
have that:

$$w_d^{(i)} = w_d^{(i-1)} + \eta x_{nd} y_n$$

where \( d \in \{1, \ldots, D\} \). Multiplying \( x_{nd} \) by \( \eta \) and \( y_n \) simply involves scaling \( x_{nd} \) and changing its sign, which can be performed by a logic gate and an inverting op amp. Moreover, deciding whether an update must be performed involves comparing \( \sum_{d=1}^{D-1} x_{mn} w_d \) to \( w_D \), where \( w_D \) is the negative threshold as before. This summation can be performed by an op amp configured as a summing amplifier. Thus, increasing the complexity of the perceptron to include more components is a straightforward extension in which multiple computing units operate simultaneously.
Navigating an environment

We now discuss a simple hypothetical paradigm in which an organism uses our hybrid analog-digital computer to make decisions about whether or not to move in the presence of multiple environmental stimuli. These stimuli may include the presence of a predator, the need to eliminate waste, or the desire to obtain food. More than one of these stimuli may be present at once, and hence, the organism must learn how to encode and process the stimuli such that it increases its chances of survival. The reactor and hardware perceptron thus function together as the organism’s brain, providing an architecture by which to encode the stimuli (reactor) and respond with or without a movement (perceptron).

As we have described in chapter 4, the reactor is a system of toroidal nanocrystalline cores that is each capable of performing a soft XOR operation. When magnetic fields of two different frequencies interact inside the Nanoperm ores in the presence of a static magnetic field, a low frequency harmonic is produced at the difference of the two input frequencies. We demonstrated that by taking an average of the spectrum of the measured flux density in the range of possible difference frequencies weighted by the magnitudes of each peak, we are able to obtain a correlation of nearly 0.9
with the true difference frequency.

We have also described in chapter 5 that we are able to build a working prototype of a perceptron in hardware. This perceptron contains a single weight which is a threshold above which the perceptron outputs high and below which it outputs low. The weight is stored as a voltage on a capacitor. Moreover, this threshold is learned as the perceptron is fed pairs of inputs and labels.

Here, we show that the intensity of the stimuli which the organism is presented with can be encoded by frequencies of currents that are magnetizing the Nanoperm cores. Moreover, the environmental parameter that dictates whether the organism should or should not move is learned by the organism and stored as its perceptron weight. Together, these two elements allow the organism to learn by experience, similar to how a real organism evaluates multiple choices and learns by trial and error.

6.1 Analog

The neuromagnetic reactor is the analog part of the organism’s “brain,” consisting of seven Nanoperm toroidal cores each of which processes a pair of stimuli. Thus, the reactor functions as the brain’s sensory processing unit. The processing is performed in layers, with three cores working together in the first layer, three in the second, and a final one in the third. The choices of which cores operate in which layers is arbitrary since there is no direct, real-time interaction among the cores (although this is a possibility for future designs). Here, we outline how the stimuli are encoded and how they are transformed through the layers to come up with a movement score.

We restrict ourselves to the three stimuli listed above and label them 1, 2, and 3 as follows:

1. drive to avoid a predator
2. drive to eliminate waste

3. drive to obtain food

Each of these stimuli independently would compel the organism to move as its intensity increases. For the sake of illustration, let us consider the case in which a linear increase in each of these independent stimuli would result in a linear increase in the desire to move. Here, we consider how the organism would deal with more than one of these stimuli being present at once. We also restrict the range of the stimuli to be integers valued and in the range 0 to 20. For example, an intensity of 0 for stimulus 1 would suggest that there is little drive to avoid a predator or that there is no predator nearby. Similarly, an intensity of 20 would suggest that a predator is very close by and threat is imminent.

For any pair of stimuli, the organism performs a soft XOR operation to calculate the degree to which it thinks it should move based on those stimuli. The intensity of a stimulus is encoded linearly as frequency, so the higher the intensity, the higher the frequency. To determine what frequencies are induced in the core, we perform an experiment in which we drive ACs with magnetic field strengths 2500 to 5000 mA/cm for two seconds to generate magnetic flux and then measure the induced flux with the measurement coil. For instance, take stimuli 1 and 2. If there is no predator nearby and there is no desire to eliminate waste, the organism does not need to execute movement of any kind. This corresponds to XOR(0,0) = 0. As shown in chapter 4, the Nanoperm cores do the best job approximating the soft XOR operation for inputs in the range 50 to 70 Hz (a difference of 20 Hz). Thus, to encode to logical zeros, we could input two 50 Hz sinusoids (one per AC coil), turn on DC, and average the frequencies of flux measured in the range 0 to 20 Hz. Looking up the result in figure 4.68c, we find that the weighted average difference frequency \( f^* \) for two 50 Hz inputs is 4 Hz. We can then perform this pairwise computation for all three pairs of
stimuli - (1, 2), (1, 3), and (2, 3) - with three cores that work in parallel in the first “layer.” Each of these would result in a difference frequency that tells us whether the organism has a desire to move.

We also outline the interpretation of the three other cases of logical inputs. Consider stimuli 1 and 2 again. If there is no predator nearby, but there is a strong desire to eliminate waste, the organism would have a strong desire to move (XOR(0,1) = 1). Similarly if there is a predator very close by and the organism also has no desire to eliminate waste, it would have a strong desire to move to evade the predator (XOR(1,0) = 1). Finally, if both the predator and desire to eliminate waste are present, the organism gets stuck in a state of decision paralysis, resulting in a low desire to move (XOR(1,1) = 0).

It is also important for the organism to make its decision in the face of uncertainty. In other words, if one pair of stimuli elicits a very strong desire to move while another pair elicits a strong desire to stay put, then the organism would have a high degree of uncertainty about its desire to move. On the other hand, if both pairs elicit the same type of desire, there is little uncertainty for the organism about what it wants to do.

To evaluate uncertainty, we compare the outputs of the first layer by feeding them in as inputs into the second. To do this, we must first perform a transformation on the outputs of the first layer so they fall into the expected range of 50 to 70 Hz. From the data plotted in figure 4.68c, we find that the minimum difference frequency is 2.8 Hz which occurs for an input of (69 Hz, 70 Hz) and the maximum difference frequency is 14.3 Hz which occurs for an input of (51 Hz, 68 Hz). To force these outputs to lie in the desired range, we can perform an affine transformation. For an input frequency $x$ in the range $[x_1, x_2]$, we can map it to $y$ in the range $[y_1, y_2]$ using
the following expression:

\[ y = (x - x_1) \frac{y_2 - y_1}{x_2 - x_1} + y_1 \]  

(6.1)

Here, \( x_1 = 2.8 \), \( x_2 = 14.3 \), \( y_1 = 50 \), and \( y_2 = 70 \). This computation allows us to compare drives elicited by all three pairs of stimuli. If two drives are very different, the uncertainty will be high. On the other hand, if the drives are same, the uncertainty will be low. (Note that 6.1 is a generic formula for affine transformation. It will also be used to convert from frequency to voltage further below.)

Finally, we combine the outputs from the first set of three cores with those of the second set by feeding them into a single core in the third layer. Since each of the first two layers produces three outputs, we average the outputs in the respective layers to determine the input from that layer into the final layer. This averaging is performed offline in software using MATLAB. The output of the final layer provides the organism with a summary of its desire to move based on the information provided by the stimuli. If the drive to move and the uncertainty are both low or both high, then the organism’s desire to move will be low. The former is straightforward and the latter is a result of decision paralysis due to the high uncertainty. On the other hand, if the drive to move is high and the uncertainty is low or vice versa, then the organism’s desire to move will be high. The former is straightforward and the latter is a result of high uncertainty.

We label the above cases I through IV and consider each of them in an investigation of the theoretical limits on the difference frequency of the output core if only the fundamental frequencies were induced in the magnetization in the input cores (i.e. no harmonics besides the difference frequency are induced). In reality, additional harmonics are induced, but because the correlation between the calculated difference frequency using the magnetic flux induced in toroids in the presence of a static magnetic field with the truth is so high, these theoretical results allow us to
make useful conclusions about the nature of our device.

We first provide some notation. We denote the frequency-encoded intensities of the three stimuli by \( \mathbf{a} = [a_1, a_2, a_3]^\top \), the differences computed by the first layer of cores by \( \mathbf{b} = [b_1, b_2, b_3]^\top = [|a_1 - a_2|, |a_1 - a_3|, |a_2 - a_3|]^\top \), and the differences of the differences computed by the second layer of cores by \( \mathbf{c} = [c_1, c_2, c_3]^\top = [|b_1 - b_2|, |b_1 - b_3|, |b_2 - b_3|]^\top \). These differences are the outputs of the first two layers, and the transformed outputs of layer 1 and layer 2 are given by \( \mathbf{b}^* \) and \( \mathbf{c}^* \), respectively. The average of the transformed outputs are denoted by \( \bar{\mathbf{b}} \) and \( \bar{\mathbf{c}} \); these are the inputs to single core in the final layer. The output of the core in the final layer is given by \( d = |\bar{\mathbf{b}}^* - \bar{\mathbf{c}}^*| \). A diagram of this is given in figure 6.1.

We can now obtain some theoretical bounds on the minimum and maximum possible drive to move and its uncertainty. First, we note that rather than using the empirical difference frequencies to transform outputs of one layer into the range appropriate for the next layer, we use the theoretical values. Thus, in formula 6.1 above, we choose \( x_1 = 0, x_2 = 20, y_1 = 50, \) and \( y_2 = 70 \). We consider extreme versions of cases I through IV above. Cases I and II correspond to the situation in which all three inputs \( \mathbf{a} \) are the same. For example, take the case \( \mathbf{a} = [55, 55, 55] \).

In this case, \( \mathbf{b} = [0, 0, 0], \mathbf{b}^* = [50, 50, 50], \mathbf{c} = [0, 0, 0], \mathbf{c}^* = [50, 50, 50], \bar{\mathbf{b}}^* = 50, \bar{\mathbf{c}}^* = 50, \) and \( d = 0 \). The low value of \( d \) is outputted regardless of the exact values in \( \mathbf{a} \), which shows that across-the-board low or high stimulus drives result in a low desire to move. The former case is straightforward to interpret while the latter can be interpreted as a result of decision paralysis due to conflicting drives.

Cases III and IV correspond to the situation in which either the average drive to move is high or the average uncertainty about the drive is high, but not both. The extreme version of case III occurs when the three input frequencies are spaced the furthest apart. This occurs for any permutation of \( \mathbf{a} = [a_1, a_2, a_3] = [50, 60, 70] \). Taking this set as our example, we find that \( \mathbf{b} = [10, 20, 10], \mathbf{b}^* = [60, 70, 60], \mathbf{c} = [20, 30, 20], \mathbf{c}^* = [70, 80, 70], \bar{\mathbf{b}}^* = 50, \bar{\mathbf{c}}^* = 50, \) and \( d = 0 \).
Figure 6.1: Block diagram showing the three layers of the neuromagnetic reactor. The first layer consists of three Nanoperm cores which are driven by pairs of sinusoidal current in the frequency range 50 to 70 Hz generating a field strength between 2500 and 5000 mA/cm. Pairwise differences of the frequencies \( a_1, a_2, \) and \( a_3 \) are approximated by averaging the spectrum in the frequency range 0 to 20 Hz, and the outputs \( b_1, b_2, \) and \( b_3 \) are generated. These are transformed in software back to the range of 50 to 70 Hz, generating the sinusoidal inputs of frequency \( b_1^a, b_2^a, \) and \( b_3^a \) to the second layer. The pairwise difference frequencies are again approximated by averaging the spectrum in the same range, producing outputs \( c_1, c_2, \) and \( c_3. \) The \( c_i \)'s are then transformed back to the range of 50 to 70 Hz in software, yielding \( c_1^a, c_2^a, \) and \( c_3^a \) (not shown). The averages of the \( b_i^a \)'s and \( c_i^a \)'s are then computed, and fed in as sinusoidal inputs to the core in the third layer. This ultimately generates an output \( d \) which is then transduced into a voltage and fed into the perceptron.

\[
c = [10, 0, 10], \quad c^a = [60, 50, 60], \quad \bar{b}^a = 63.333, \quad \bar{c}^a = 56.667, \quad \text{and} \quad d = 6.667.
\]

Thus, we find that when the average drive to move is low, the difference is high, signaling the organism to move.

Case IV occurs when the drive to move is low but the uncertainty is high. Graphically, in figure 6.2, we find that \( \bar{b}^a > \bar{c}^a \) for all choices of \( a. \) This means that the maximum uncertainty scales with the value of \( \bar{b}^a. \) The maximum difference of 6.667 between \( \bar{b}^a \) and \( \bar{c}^a \) occurs when the three elements of \( a \) are spaced the furthest apart. Moreover, from cases I and II as well as from figure 6.2, we know that the smallest difference of 0 occurs when the three elements of \( a \) are equal. Thus, case IV
Figure 6.2: $\bar{b}^*$ and corresponding $\bar{c}^*$ values for all possible values of elements $a$.

essentially encompasses all other combinations of the entries of $a$. Graphically, this includes all indices besides the 6 on the very left or the 21 on the very right of figure 6.2. These correspond to the six permutations of the elements of $a$ and the 21 ways in which the values of $a$ can all be equal by all taking on values between 50 and 70 (inclusive).

In practice, the difference frequencies computed by the cores are not exactly the difference of the two inputs; they are weighted averages of the spectrum of the induced flux from 0 Hz to 20 Hz. Thus, the practical results differ slightly from theory.

Thus, the reactor allows the organism to evaluate the stimuli it is presented with through their magnitude and their uncertainty. It encodes the drive provided by each of the stimuli as a frequency between 50 and 70 Hz and uses a set of parallel nonlinear transformations to capture the overall desire of the organism to move. Once the final difference frequency is computed, the value is scaled into one last time using formula 6.1. This value is treated as a voltage which is an input to the
perceptron. Together with a label, the perceptron uses these voltages to learn an environmental parameter that determines what decision the organism should make when faced with any combination of the three stimuli.

6.2 Digital

In this section, we describe the digital part of our organism’s “brain,” which is the perceptron discussed in chapter 5. Our hardware perceptron consists of a single parameter which is a threshold above which the output is classified as a +1 and below which the output is −1. In our hypothetical environment, +1 would correspond to the decision to move while −1 would correspond to the decision not to move. Thus, the perceptron represents the motor output part of the organism’s brain. Previously, we described circuit simulation results demonstrating that our perceptron design was capable of implementing the perceptron learning algorithm (PLA) and learning a weight from the labeled input data. Here, we show experimental results that demonstrate that the physical hardware is also capable of learning the weight as expected based on the simulation.

The data that the perceptron classifies are obtained by transducing the frequency measured/computed at the output layer of the reactor into a voltage. Recall that this frequency \( d \) is a weighted average of the spectrum in the range 0 to 6.667 Hz. We convert this frequency to a voltage \( d^* = x \) in the range \([-0.4, 0.4]\) using equation 6.1 by choosing \( x_1 = 0, x_2 = 6.667, y_1 = -0.4, \) and \( y_2 = 0.4 \). This range was chosen arbitrarily, and we are only limited by the supply voltage of the op amps. Specifically, the doubling operation performed during the compute phase of the perceptron doubles the inverted sum of its inputs. Because the op amps have a supply voltage of 5 V, the input must be less than 2.5 V, so values whose magnitudes are less than 2.5 V are acceptable. However, we also note that the output of the reactor is discrete since (1) we restricted our inputs to be integer frequencies so the
output, which is a result of computing a finite number of absolute differences and averages of integers, must also be discrete and (2) repeating a set of inputs in separate experimental trials yields the same spectrum and the average frequencies each time. By choosing a smaller range of voltages over which the frequencies outputted by the reactor are transduced, we reduce the gap between neighboring discrete values, allowing us to demonstrate that the perceptron is capable of learning robustly even when the gap between points that are labeled \(-1\) and those that are labeled \(+1\) is as small as 80 mV.

The perceptron is a supervised learning algorithm, and hence, learning can only take place on labeled data points. In our setting, the labels are generated on the fly using the following rule:

\[ y = \begin{cases} 
-1, & x < -0.2. \\
1, & \text{otherwise.}
\end{cases} \]

The choice of this threshold is arbitrary. We chose -0.2 because it is nontrivial (not zero) and falls halfway between 0 and one of the bounds (-0.4). This corresponds to a difference frequency \(d\) of 1.667 Hz.

6.3 Hybrid

We now give experimental results showing how the analog reactor and digital perceptron work together to learn the threshold defined above. We remind the reader that this was a proof of concept, so much of the processing was performed offline. We also use the vectors \(a\), \(b\), and \(c\) somewhat loosely to refer to not only the fundamental frequencies but also their harmonics.

We started by generating a series of 180 trials comprised of a triplet of stimulus frequencies which we fed into the first layer of the reactor at a magnetic field strength between 2500 and 5000 mA/cm. These were inputted pairwise as shown in figure 6.1.
Figure 6.3: Spectrograms for the three Nanoperm cores in the first layer of the reactor showing the input frequencies \( a \) (left subplots, range of 50-70 Hz) and induced difference frequencies and harmonics (right subplots, range of 0-20 Hz) for each of the 180 trials. Colors indicate the strength of the induced flux density in mT.

Figure 6.3 shows the spectrogram of stimuli for the first layer of the reactor. Peaks in the left subplots indicate the input frequencies \( a \) while peaks in the right subplot indicate the difference frequencies and their harmonics. For each trial, we performed a weighted average of the power spectrum in the range 0 to 20 Hz using equation 4.9 and the resulting difference frequencies \( b \) were transformed into the range 50 to 70 Hz using equation 6.1.
These \( b_i^* \)'s were then fed into the second layer consisting of three cores. Figure 6.4 shows the analogous spectrograms for the second layer of processing, where the inputs are the \( b_i^* \)'s. Processing performed by the second layer results in another set of difference frequencies \( c \) which fall in the range 0 to 20 Hz. These are then transformed into the range 50 to 70 Hz, yielding \( c^* \). The \( b_i^* \)'s and \( c_i^* \)'s are then averaged in software to get \( \bar{b}^* \) and \( \bar{c}^* \).
Figure 6.5: Spectrograms for the final Nanoperm cores in the third layer of the reactor showing the input frequencies $\tilde{b}^s$ and $\tilde{c}^s$ (left subplots, range of 50-70 Hz but restricted mathematically to 50-63 Hz) and induced difference frequencies and harmonics (right subplots, range of 0-20 Hz but restricted mathematically to 0 to 7 Hz) for each of the 180 trials. Colors indicate the strength of the induced flux density in mT.

These are then fed in as inputs to the core in the third layer where the final difference frequency $d$ is computed. Figure 6.5 shows the analogous spectrograms for the third layer of processing. Note that the major peaks in the left subplot (which contains the inputs $\tilde{b}^s$ and $\tilde{c}^s$) all fall below 63 Hz. This is because the largest value the average of the three inputs can take on is 63.333 Hz because of the restriction that the same three input frequencies are shared among the three pairwise comparisons. Thus, to maximize the average of the absolute differences, two must equal 70 Hz while the third is 50 Hz. Also notice that the major peaks in the right subplot (which is used to compute the output $d$) fall below 7 Hz. This is also because of the same restriction and was discussed in section 6.1. Thus, our experimental results on the reactor agree with our expectations.

With the spectrum of the output of the reactor, we can now compute $d$ using
equation 4.9 and transduce it into a voltage $x$ that is inputted into the perceptron with equation 6.1. We perform the process of generating synthetic stimuli, running them through the reactor, and computing $d$ for all 180 trials in batch mode offline. Then, using the values of $d$, we generate a new time series which we feed into the perceptron (along with the labels).

Figure 6.6 shows a snapshot of the first 300 ms of $x$ overlaid on $V_{\text{timing}}$, the voltage across the capacitor in the timing circuit. As expected, we find that $x$ changes its value whenever the $V_{\text{timing}}$ is increasing and crosses 2.0 V. We accomplished this synchronization between $V_{\text{timing}}$ and $x$ by first running the perceptron for ten seconds; since the time constant of the timing circuit is about 4.7 ms and five time constants are required to completely charge and discharge the capacitor, respectively, we approximated the total session duration to be $(180 \text{ trials}) (4.7 \times 10^{-3} \text{ sec/time constant} \times 10 \frac{\text{time constants}}{\text{trial}}) = 8.5 \text{ sec}$, so 10 seconds of recording time was more than enough. During this run, no inputs were transmitted; we simply recorded the $V_{\text{timing}}$ to get
Figure 6.7: Plot of the values of $d$ corresponding to $x$ from figure 6.7. The values from the first several trials for the reactor experiment are listed in table 6.1. By inspection, we see that they match the values of $d$ plotted here.

a realistic picture of how the voltage across the timing capacitor behaves with the Schmitt trigger. Then, using the measured time series for $V_{\text{timing}}$, we created input vectors $x$ and $y$. The values of $x$ correspond to those obtained from the preceding reactor experiments. To see this, we plot $d$ (the frequency corresponding to the voltage $x$) for the first 300 ms of the new time series in figure 6.7 and list out the difference frequencies obtained by the reactor experiment for the same time window in table 6.1. We see by inspection that the frequencies in the plot match those in the table. Finally, comparing figure 6.6 to figure 6.7, we see that $x$ robustly encodes $d$ after the transformation by equation 6.1.

The time series for $y$ is generated in a similar way except it takes on only two values: either $+5$ V when $x > -0.2$ or $+0.5$ V when $x < -0.2$. These two correspond to logical \texttt{high} and \texttt{low}, respectively. (The logic gates used in the perceptron accept voltages greater than 3.85 V as logical \texttt{high} and voltages less than 1.65 V as logical \texttt{low}, so any voltage that falls in these ranges is appropriate.) Figure 6.8 shows a plot
Table 6.1: Resulting difference frequency $d$ for the first ten trials of the reactor experiment. Comparing these values to those plotted in figure 6.7 shows that the time series $x$ generated for the perceptron by processing the stimuli in batch mode takes on values that match those computed by the reactor.

<table>
<thead>
<tr>
<th>Trial</th>
<th>$d$ (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.667</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>2.667</td>
</tr>
<tr>
<td>4</td>
<td>0.667</td>
</tr>
<tr>
<td>5</td>
<td>6.667</td>
</tr>
<tr>
<td>6</td>
<td>1.333</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0.667</td>
</tr>
<tr>
<td>9</td>
<td>0.667</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
</tr>
</tbody>
</table>

of $y$ for the first 300 ms. We see that the labels for $x$ are correctly encoded by $y$.

Having created time series for $x$ and $y$ in correspondence with our measured $V_{\text{timing}}$, we need to utilize this time series to run the perceptron. Rather than running the perceptron off the timing circuitry used to generate $V_{\text{timing}}$, we directly input the three time series $V_{\text{timing}}$, $x$, and $y$ into the perceptron. This way, we ensure that the timing of $x$ and $y$ are in sync with the rest of the operations performed by the perceptron. In future designs, we can eliminate this by utilizing sample-and-hold circuitry. In summary, this circuit would sample a value outputted from the NI-DAQ and maintain this value until a threshold-crossing condition is met.

Next, we depict the results of the phase identifier. As in figure 5.14, when we record the output of the window comparators, we expect them to go high when each of their windows is triggered. Figure 6.9 shows the voltage at the output of each of the window comparators overlaid on a single charge-discharge cycle of $C_{\text{timing}}$. The solid line corresponds to reset1 activating, the dashed line to compute, the dotted line to reset2, and the dash-dotted line to update. Each of the windows appears twice - once when $C_{\text{timing}}$ is charging and again when it is discharging.
Figure 6.8: Plot of the time series of $y$ generated by first recording $V_{\text{timing}}$ and then creating a vector whose values change when $V_{\text{timing}}$ is rising and crosses 2 V. The values of $y$ correctly label the corresponding $x$ values plotted in figure 6.6.

This is expected since the outputs of the window comparators are independent of the charging / discharging state of $C_{\text{timing}}$. When we look at the output of the AND gates to which the output of the window comparators are connected (shown in figure 6.10), we find that the four windows that are present when $V_{\text{timing}}$ is charging have disappeared. This is because the high-pass filter / differentiator outputs high only when $V_{\text{timing}}$ is discharging, and the outputs of the AND gates require both the window comparator and the differentiator to be high. Note that we had to shift to a different part of the experiment to report this result because the first incorrect prediction was not made until nearly 500 ms had elapsed; the prediction being incorrect is the third and final condition to trigger the AND gates. We also find that while the windows are all appropriately triggered when looking at the output before the AND gate (figure 6.9), but after the AND gate (figure 6.10), the duration of the update phase is cut short. Moreover, an artifact appears in the voltage trace for both $v_{\text{timing-decreasing}}$ and Inc between the reset2 and update phases.
Despite thorough investigation, we were unable to find an explanation for this but found that it did not significantly affect the performance of the hardware.

Finally, we show that the hardware perceptron learns a weight to appropriately discriminate between data that lies in class +1 and data that lies in class −1. We omit here the detailed experimental results as they are quite similar to the simulated results. To summarize, we show in figure 6.11 that the voltage on the weight-storing capacitor $w$ rapidly converges to approximately 0.2 V. Recall that the threshold we arbitrarily set to label our data was -0.2. Thus, the value learned by the perceptron makes sense since, as explained previously, the weight it learns should be the negative threshold. This value slowly decays due to leakage currents that cause the capacitor to discharge. When the voltage falls below approximately 0.15, the perceptron makes an error in its prediction and the voltage on the weight storing capacitor is updated so that it again hovers around 0.2 V. This process is depicted in figure 6.11.

We have demonstrated that the perceptron successfully learns the environmental
Figure 6.10: Plot of the output of the logic gates to which the phase identifier window comparators are connected for a single charge-discharge cycle of $C_{\text{timing}}$ when an incorrect prediction is made. Note that we had to shift to a different part of the time series to report this result because the first incorrect prediction was not made until nearly 500 ms had elapsed. While the first three phases $\text{reset1}$, $\text{compute}$, and $\text{reset2}$ are triggered correctly, the final phase $\text{update}$ is cut short. Moreover, an artifact appears in the voltage trace for both $\text{vtiming-increasing}$ and $\text{Inc}$ between the $\text{reset2}$ and $\text{update}$ phases. Despite investigating, we were unable to find an explanation for either of these; however, the performance of the circuit did not suffer.

Parameter -0.2 based on the labeled inputs that were processed by the reactor. This information allows the hypothetical organism to make better decisions in the face of uncertainty as it learns by trial and error. While our hybrid analog-digital hardware is by no means perfect, it demonstrates as a proof-of-concept how to utilize signals that are encoded as magnetic fields to perform nonlinear computations that allow for learning from data. Future iterations of the design could utilize feedback between the analog and digital subsystems to enable a recurrent system. Such a system would be even more adept at processing time series data and would allow for more robust memory storage. Moreover, the hardware perceptron could also be improved to allow for learning of more parameters. The system could be built up hierarchically to function similarly to a hardware neural network.
Figure 6.11: Plot showing the voltage $V_w$ on the weight-storing capacitor. The values of $x$ are plotted as colored circles rather than as a continuous trace to highlight the label of the data points. Note that the weight of approximately 0.2 achieved by the perceptron is close to the negative of the threshold -0.2, showing that our perceptron is successful at learning the environmental parameter. Anomalies exist in the $V_w$ time series. For instance, around 2.5 sec, we see the voltage slowly start to rise. This could be due to a mistiming of the learning circuitry, which, as we described in chapter 5, may contain photorelays that are being operated outside their optimal on-time and off-times. Second, around 4 sec, we see that learning takes place once successfully but then immediately happens again, resulting in an abnormal slow rise in $V_w$. The explanation for this may also have to do with the suboptimal timing of the photorelays but remains to be confirmed.
In summary, we first demonstrated the plethora of sources of magnetic fields generated by neural tissue and how they can be used for signaling and diagnosing medical conditions. We then showed examples of how the brain, consisting of neurons which generate these magnetic fields, has inspired two major categories of brain-inspired computing: neural networks and neuromorphic hardware. Neural networks with backpropagation have been implemented in physical media through acoustic, electro-optical, and purely optical methods, bridging the two categories. In addition, the observation that the brain is a hybrid analog-digital computer that cannot be simulated by a Turing machine inspired Cicurel and Nicolelis to develop a theory known as the relativistic brain theory (Cicurel and Nicolelis, 2015). According to this theory, neural activity and the electromagnetic fields that are induced interact to perform computations and store information, categorizing the brain as hybrid analog-digital. Based on this theory, we were inspired to create a hybrid computing device that could generate nonlinear features in analog hardware through interactions of magnetic fields and then employ these features in digital hardware to perform nonlinear pattern recognition. Our device is a proof-of-concept but with the appropriate
innovations has a number of practical applications, both in terms of understanding brain function as well as in the field of hybrid analog-digital computing.

The 3D printed design of the reactor could serve as a plausible first step to building a biologically realistic brain model in hardware. In this model, the tract shapes are directly taken from white matter tractography; as such, it would be the first hardware of its kind to directly incorporate the geometry of tracts. By passing currents through these tracts and measuring the induced magnetic fields, we can get a sense of how the distribution of magnetic flux changes over space and time within the brain. According to the RBT, the major white matter pathways contribute significantly to endogenous neuromagnetic fields and may be responsible for inducing psychiatric conditions such as autism spectrum disorder and schizophrenia. By studying how the distribution in flux density changes when the shape of the white matter is altered, we can quantify the role that white matter configuration plays in neural computation.

However, this would require some major changes to the design of the device. First, the currents that are generated would have to be scaled down so that they are on the order of magnitude of the currents in the brain; this would require a reduction in current by about five to ten orders of magnitude, depending on the number of neurons that are in a single bundle. The reduction in current amplitude would also reduce the extent to which the electromagnetic fields induced by the straight tracts have an effect. To account for possible amplification in the field strength, we would need to factor in the smaller loops of white matter that are present throughout the brain. These might serve as transformers, which are capable of amplifying magnetic field strength based on the number of turns. Thus, rather than utilizing a global, anatomically constrained tractography for the design, we might utilize a local tractography. This would result in a 3D printout such as the one shown in figure 7.1.

Second, propagation of currents down the tracts in the brain is much slower (ap-
Figure 7.1: Local tractography analysis and 3D printout. A model based on this (perhaps with even finer resolution) could be used to realistically model the white matter pathways in the brain. More loops of white matter are identified than in the global, anatomically constrained tractography. By incorporating Hall effect sensors to model neurons, one could also study how neurons and magnetic fields interact.

approximately 50 m/s) than that of wires in the model (approximately $3 \times 10^8$ m/s). To slow down the propagation of signals, we could employ a series of transistors. These transistors would gate the incoming current, and when the gating signal reaches a threshold, the source and drain conduct allowing the signal to pass. The gating signal might be clocked to ensure that the current is limited to a certain propagation velocity. The use of transistors mimics the approach taken by more traditional neuromorphic devices. Another minor improvement could be made by incorporating transmembrane currents into the model; while these were shown by (Swinney and Wikswo Jr, 1980) not to be significant in their contribution to the macroscopic magnetic field, they nevertheless play a small role.

Third, the magnetic properties of the medium should reflect those of the brain. The brain is a conductive tissue supplied by vasculature containing iron particles.
Oxygenated blood is weakly diamagnetic (weakly repels or cancels external magnetic fields) while deoxygenated blood is weakly paramagnetic (is weakly attracted by external magnetic fields). Aside from these particles, the brain can be approximated as an aqueous bath of ions; these are not magnetic themselves but, in theory, allow for very weak inductive effects to take place.

These changes might also lead to useful computational advantages. For instance, another major addition that could be performed is the feedback between neuronal action potentials and induced fields. This could be done by extending the 3D array of Hall effect sensors employed in the second iteration of our device. These Hall effect sensors could imitate neuronal cell bodies, picking up magnetic fields induced by the axons and transducing them into a voltage. The voltage could be amplified and directly fed back into the tracts. In this way, we would form a closed loop system between the tracts representing axons and the Hall effect sensors representing cell bodies.

This system would then behave as an extension of that described in (Hermans et al., 2015). Hermans showed that one could generate very high dimensional feature spaces from low dimensional inputs by encoding individual samples as sounds that were fed through an acoustic medium consisting of nonlinear feedback. The signals could then be measured with microphones and converted back into discrete samples, and a prediction error could be calculated and backpropagated through the system by simply inverting the locations of the sources and receivers.

To demonstrate feasibility, they employed a system consisting of a single speaker and a microphone placed inside a hollow tube. Sounds played inside the tube by the speaker are delayed and induce harmonics that are picked up by the microphone. The recorded signals are then fed back in through the speakers. The error between the prediction and the encoding of the true output is then calculated on a computer and is backpropagated to determine the gradients that are necessary to optimize the
system parameters. A similar setup is employed for optical systems.

In our system, the sources would initially be the currents passed through the tracts. These would generate magnetic fields which pass through the conductive medium. The induced flux densities could be measured by the array of Hall effect sensors, and then, as in (Hermans et al., 2015), a rectified linear (ReLU) transformation implemented in hardware could be applied before feeding the signals back into the tracts as voltages. Neurons are known to exhibit smooth relaxations of ReLU nonlinearities, making this a plausible model for feedback between biomagnetic fields and the currents that generate them. As mentioned in chapter 5, the derivative of the ReLU function is either 0 or 1 depending on whether the input is less than or greater than 0. This makes the derivative easy to compute analytically and implement in hardware.

A related idea involving feedback is depicted in figure 7.2, which came from one of our preliminary computational studies. In this approach, we imported the tractography into MATLAB and simulated an interaction between neuronal activity and induced fields. Spikes were allowed to propagate down the tracts and field effects were modeled as spatially decaying Gaussians whose amplitude is given by the voltage along the tract. Fields extended spatially in two directions along the axis of the tract. (In reality, fields extend in all three dimensions and do not decay exponentially, but this simulation was designed for computational purposes rather than for investigating the brain’s electromagnetic fields.) The induced fields could trigger additional action potentials when the sum of the field strength at a given point along the axon exceeded a threshold. We showed that spontaneous spikes that were used to initialize the system generated self-sustaining oscillatory behavior as a result of feedback from the fields. With the appropriate choice of decay functions and by allowing the spike waveforms to adapt, it may be possible encode information in these oscillations. This would represent a discrete approximation to a true hybrid
(a) Simulated spikes propagating down tracts.  (b) Induced voltages due to the spikes.

Figure 7.2: Simulation of spikes and induced magnetic fields interacting and encoding information by oscillations. The tracts were directly imported from the tractography analysis. 2D Gaussian functions were chosen to model the spatial decay of the induced field due to a voltage. When the induced field at a point along an axon exceeds a threshold, an additional spike is induced at that location. Over time, this sets up an oscillatory behavior, despite the input being only a spontaneous impulse of action potentials on a subset of the tracts. This oscillation can be used to encode information.

Note that the above also provides an additional reason to utilize the DTI to inform the shape of the reactor. Detecting differences between the magnetic field patterns induced in healthy brains and those of subjects with ASD, for example, would inform us about how to introduce magnetic stimulation to correct for the abnormalities. For instance, one could employ the micromagnetic stimulators described in (Bonmassar et al., 2012). These coils are only 1 mm in diameter, 1 mm long, and contain about 21 turns; they generate electric fields that permeate the space about 200 µm from the edge of the coil and have an amplitude of 6 V/m. Hence, we could in theory
determine where to place these sensors and how to orient them. One challenge with these stimulators, however is that they employ large currents on the order of 10 A and can only be used to carry 5 \( \mu s \) pulsed DC currents before melting occurs. This is an issue that still needs to be corrected in the stimulator design if it is to be used in vivo.

Returning to the computational utility of the device, we propose two major extensions that both incorporate nonlinear feedback. We consider here the Nanoperm toroidal core design (version 4) of the reactor for simplicity but the ideas can be generalized to the other designs. First, we show that we can introduce memory into the device by incorporating feedback between two cores in the following way. Recall that each core consists of three inputs - two AC and one DC - and one output. One input to each of the cores is an independent input; it is not involved in feedback. The second input is the filtered, amplified measurement of the induced flux on the other core. Initially, one input to one core is turned on. The flux measured on that core is filtered, amplified, and fed into the second core as an input, along with an additional independent input. These two are then combined (nonlinearly), and the induced flux is filtered, amplified, and fed back into the first core. Through the measurement of this induced flux density, the first core indirectly sees the independent signals applied to both cores. This total signal is now filtered, amplified, and passed back to the second core. After some time, the measurement on both cores will equilibrate to be the (nonlinear) sum of the flux induced on each of the cores. The independent inputs are no longer needed and thus can be switched off. Assuming proper gains are chosen for the feedback amplification, the signal still propagating between the cores should now represent the (nonlinear) sum of the inputs to each of the individual cores. Thus, the system is able to memorize the sum of the inputs applied to it since it no longer needs those inputs to generate the feedback signal. A diagram of this circuit is given in figure 7.3.
Figure 7.3: Memory circuit in which the measurements made on each of the toroids are filtered and fed back to each other. Because the independent inputs to each of the toroids can be turned off after the system has reached equilibrium while the sum of the inputs and their harmonics continue to propagate between the cores, we propose that this device exhibits memory.

Second, the device can be used a simulator for a chaotic system. Chaotic systems require three or more dynamic quantities to interact with nonlinear feedback. If only two are present, the system can converge to an attractor, oscillate periodically, or diverge; however, if three or more interact, the system can suddenly escape from an orbit that almost looks periodic and return to a quasi-periodic state. This is known as chaos.

Chaotic systems are highly sensitive to initial conditions. Slight differences can cause the outputs after some amount of time to diverge. They are known to exist in the brain based on studies of neural circuits. It is possible that chaotic interactions could also be induced by feedback among magnetic fields induced by neural activity.
and/or the interaction between magnetic fields and neural circuitry. According to
the RBT, it is possible that this chaotic nature gives rise to phenomena such as
creativity and consciousness. By utilizing the reactor as a simulator of chaos, one
could start to develop ways to test these hypotheses.

To implement this, we would simply extend the system of two cores with feedback
to a system of three or more. The simplest design would involve the first core feeding
its measurement to the second; the second to the third; and so on. The last core
would feed its measurement back to the first. It is important to choose the parameters
of a system such that the system behaves chaotically. In our case, the parameters
include the cutoff frequencies of the filters and the gains of the amplifiers. If these are
not in the appropriate range, the system will simply converge, oscillate, or diverge.
A diagram of the circuit is given for the three-core case in figure 7.4.

From a technological standpoint, one might also be able to develop new kinds
of memristive technology using this device. Recall that memristors are circuit el-
ements whose resistance is proportional to the history of charge that has passed
through them. This implies that the current-voltage characteristics of memristors
exhibit hysteresis loops. We found in our experiments that the applied magnetic
field strength and induced flux density in the Nanoperm cores also had a hysteresis.
By integrating the induced voltage in hardware using an active op-amp integrator
and then feeding the output into a transconductance amplifier, we can also produce
a current source which has a nonlinear, hysteretic relationship with the input volt-
age. This constitutes one possible implementation of a memristor. (Memristors can
be implemented in many ways; the exact function that the memristor integrates is
based on its design.) In addition, (Di Ventra et al., 2009) provided theoretical results
for the existence of memcapacitors and meminductors and suggested that they may
be useful for neuromorphic computing; as shown in (Pershin and Di Ventra, 2012),
this results in much simpler implementations of Boolean logic operations than with
Figure 7.4: Circuit in which three cores interact with nonlinear feedback, possibly resulting in chaotic behavior. The filter characteristics and the gains of the amplifiers must be chosen such that the system behaves chaotically.
memristive circuits alone.

As a hybrid analog-digital device that utilizes magnetic fields, our device could offer advantages over traditional neuromorphic architectures. In fact, one group has recently adopted magnetic fields in some of their latest designs of neuromorphic chips (Schneider et al., 2018). Their device utilizes niobium superconductors to model axons and magnetic manganese nanoclusters to model synapses. Thus, their device is highly power-efficient and capable of encoding information both electrically and magnetically. However, their device is still limited to only a handful of synapses per electrode while biological neurons contain hundreds if not thousands of synapses each. Moreover, the device must be cooled to near absolute zero in order for the superconductors to function.

These issues might be addressed by utilizing more traditional neuromorphic chips whose neurons can be synchronized via magnetic feedback, similar to the modified reactor model with Hall effect sensors described above. The chips could be designed in alternating layers of electrically conductive and magnetic material. The application of magnetic fields in the magnetic layers would cause the currents in the electrical layer to synchronize through magnetic induction, functioning like a magnetic ephaptic effect. The geometry of the tracts in the brain might inform us of how to constrain the magnetic fields such that they affect only pools of neurons while the network of synapses allow the pools to communicate synaptically.

This summarizes a list of possible improvements that could be made to the device, which is a hybrid analog-digital computer inspired by the RBT (Cicurel and Nicolelis, 2015). By incorporating these changes, we can develop a practical computer in which the hardware implements both the memory and instructions. Unlike traditional architectures, the device would be optimized to solve problems in which (nonlinear) dynamics govern the evolution of a system (e.g. protein folding, simulations of the stock market). For these problems, an architecture which adapts as the computation
proceeds is not only a natural choice but is also more brain-like and efficient. Our device could also be used to study the brain itself, including the ways in which abnormal white matter geometries might cause phenotypes such as schizophrenia or ASD. Thus, with the proposed modifications, our device may be effective not only as computational hardware but also as a way to explore a new paradigm for studying the brain.
Appendix A

Magnetic field and hysteresis measurements
Figure A.1: 51 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varied field strengths.
Figure A.2: 67 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at various field strengths.
Figure A.3: 95 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.
Figure A.4: 102 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at various field strengths.
Figure A.5: 127 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.
Figure A.6: 147 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at various field strengths.
Figure A.7: 150 Hz hysteresis loops (left) and FFTs of induced magnetic fields (right) in the presence of direct magnetic field at varies field strengths.
Figure A.8: Nanoperm hysteresis loops measured when driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 59.8 mA/cm.
Figure A.9: FFT of magnetic field when core is driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 59.8 mA/cm.
Figure A.10: Nanoperm hysteresis loops measured when driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 299.1 mA/cm.
Figure A.11: FFT of magnetic field when core is driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 299.1 mA/cm.
Figure A.12: Nanoperm hysteresis loops measured when driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 598.2 mA/cm.
Figure A.13: FFT of magnetic field when core is driven by two frequencies in the presence of static magnetic field induced by magnetic field strength of 598.2 mA/cm.
Bibliography


246


256


*CD74AC86 Quadruple 2-input Exclusive-OR Gate*. Texas Instruments, 1 2003.


Biography

Vivek Anand Subramanian received a B.S.E. in Biomedical and Electrical and Computer Engineering in 2012, an M.S. in Statistical Science in 2015, and a Ph.D. in Biomedical Engineering in 2018, all from Duke University in Durham, NC. After developing a deeper expertise in the fields of machine learning and artificial intelligence (AI), he hopes to expand fundamental knowledge in these fields and to develop AI-based technology that has practical impact on quality of life. When he is not performing research on neuronal time series or brain-inspired electronics, he enjoys traveling, playing video games and video game music, and watching classic and animated films.