NANOMAGNET LOGIC: FROM DEVICES TO ARCHITECTURES

A Dissertation

Submitted to the Graduate School
of the University of Notre Dame
in Partial Fulfillment of the Requirements
for the Degree of

Doctor of Philosophy

by
Aaron J. W. Dingler

Michael T. Niemier, Director

Graduate Program in Computer Science and Engineering
Notre Dame, Indiana
July 2013
NANOMAGNET LOGIC: FROM DEVICES TO ARCHITECTURES

Abstract

by

Aaron J. W. Dingler

Nanomagnet logic (NML) is an emerging device technology that exploits the magnetic state of nanoscale magnets to encode logic values. NML is being considered as a potential replacement for current charge-based technologies (namely, CMOS transistors) in order to achieve increased performance in terms of energy, delay, and/or area. The potential advantages of NML are: devices are non-volatile and radiation hard, device switching events consume little power, devices are scalable to the superparamagnetic limit, and there is a clear path to implementation that leverages current CMOS fabrication techniques. The goal of this work is to advance the state of the art for NML in order to realize these potential advantages at the application and architecture levels.

In order to evaluate the ultimate performance of NML, methods are needed to bridge the gaps between the device and circuit levels, and between the circuit and architecture and application levels. This dissertation thus considers methods to analyze NML device-to-device interactions that allow for the design of complex circuits. Next, this work considers the interaction between NML circuits and the clock required to re-evaluate the circuits. Specifically considered is how to model a proposed clock structure to determine its effects on NML circuit logical correctness, power, and energy requirements. Furthermore, this dissertation shows how uni-directional dataflow can be accomplished in the context of these realistic clock structures. This study of NML clocking allows for three studies at the architecture and application levels. First, this work shows how even with the energy
consumed in the proposed clock, NML could outperform CMOS in terms of the energy delay product for a ripple carry adder benchmark. Second, this dissertation demonstrates how to realize non-volatility in NML in order to enable non-volatile NML shift registers with low energy operation. Finally, this work demonstrates how NML could fit into a stochastic computing architecture to provide performance improvements over conventional and stochastic CMOS equivalents.
To my wife.
CONTENTS

FIGURES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . vii

TABLES . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xiv

ACKNOWLEDGMENTS . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . xv

CHAPTER 1: INTRODUCTION . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
  1.1 Overview . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 1
  1.2 My Contributions . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
    1.2.1 NML Circuits . . . . . . . . . . . . . . . . . . . . . . . . . . . . 2
    1.2.2 Modeling Realistic Clock Wires . . . . . . . . . . . . . . . . . . 3
    1.2.3 Uni-directional Dataflow, Energy Analysis, and Comparison to CMOS 4
    1.2.4 NML Shift Registers and Non-volatility . . . . . . . . . . . . . . 4
    1.2.5 Stochastic Computing . . . . . . . . . . . . . . . . . . . . . . . 5
  1.3 Organization of Document . . . . . . . . . . . . . . . . . . . . . . . . . 6

CHAPTER 2: BACKGROUND . . . . . . . . . . . . . . . . . . . . . . . . . . . . 7
  2.1 Fundamental Constructs . . . . . . . . . . . . . . . . . . . . . . . . . . 7
  2.2 Clocking . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 8
  2.3 I/O . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 11
  2.4 Tenets of Digital Logic . . . . . . . . . . . . . . . . . . . . . . . . . . . 12
    2.4.1 Functionally Complete Logic Set . . . . . . . . . . . . . . . . . . 12
    2.4.2 Power Amplification . . . . . . . . . . . . . . . . . . . . . . . . . 13
    2.4.3 Non-linear Response . . . . . . . . . . . . . . . . . . . . . . . . . 13
    2.4.4 Concatenability . . . . . . . . . . . . . . . . . . . . . . . . . . . 14
    2.4.5 Uni-directional Dataflow . . . . . . . . . . . . . . . . . . . . . . 14

CHAPTER 3: FROM NML DEVICES TO CIRCUITS . . . . . . . . . . . . . . . . . 17
  3.1 Analyzing Magnetic Circuit Elements . . . . . . . . . . . . . . . . . . . . 17
    3.1.1 Magnetization State - Traditional MH Curves . . . . . . . . . . . . 17
    3.1.2 Magnetization State - NML Representation . . . . . . . . . . . . . 18
  3.2 Shape, Size, and Circuits . . . . . . . . . . . . . . . . . . . . . . . . . . 23
  3.3 Clock Wire Boundaries . . . . . . . . . . . . . . . . . . . . . . . . . . . 28
  3.4 Horizontal and Vertical Dataflow . . . . . . . . . . . . . . . . . . . . . . 34
  3.5 Wire Crossings . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 38
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.6 Path to Lower Power: New Magnetic Materials</td>
<td>103</td>
</tr>
<tr>
<td>5.6.1 Initial Work</td>
<td>104</td>
</tr>
<tr>
<td>5.6.2 Related Work</td>
<td>104</td>
</tr>
<tr>
<td>5.6.3 Enabling Signal Propagation</td>
<td>105</td>
</tr>
<tr>
<td>5.6.3.1 Switching in the Presence of an Applied Field</td>
<td>105</td>
</tr>
<tr>
<td>5.6.3.2 Switching in the Absence of an Applied Field</td>
<td>109</td>
</tr>
<tr>
<td>5.6.3.3 Varying K1 Value</td>
<td>111</td>
</tr>
<tr>
<td>5.6.3.4 Supermalloy</td>
<td>114</td>
</tr>
<tr>
<td>5.6.4 Summary</td>
<td>117</td>
</tr>
<tr>
<td>5.7 Conclusions</td>
<td>117</td>
</tr>
<tr>
<td>CHAPTER 6: NML SHIFT REGISTER</td>
<td>119</td>
</tr>
<tr>
<td>6.1 Initial Work</td>
<td>120</td>
</tr>
<tr>
<td>6.2 Volatility in NML Ensembles</td>
<td>121</td>
</tr>
<tr>
<td>6.2.1 Related Work</td>
<td>122</td>
</tr>
<tr>
<td>6.2.2 Impact of Bit Conflicts in the Steady State</td>
<td>122</td>
</tr>
<tr>
<td>6.3 Improving Stability</td>
<td>124</td>
</tr>
<tr>
<td>6.3.1 All-Magnet Latch</td>
<td>126</td>
</tr>
<tr>
<td>6.3.2 Single-Magnet, Three-Magnet Latches</td>
<td>127</td>
</tr>
<tr>
<td>6.4 Quantifying Stability</td>
<td>128</td>
</tr>
<tr>
<td>6.4.1 Ensemble Stability</td>
<td>129</td>
</tr>
<tr>
<td>6.4.2 Examples</td>
<td>130</td>
</tr>
<tr>
<td>6.4.2.1 No Latching</td>
<td>130</td>
</tr>
<tr>
<td>6.4.2.2 All-Magnet Latch</td>
<td>132</td>
</tr>
<tr>
<td>6.4.2.3 Single-Magnet Latch</td>
<td>133</td>
</tr>
<tr>
<td>6.4.2.4 Summary</td>
<td>135</td>
</tr>
<tr>
<td>6.5 Stability Trends</td>
<td>137</td>
</tr>
<tr>
<td>6.5.1 Methodology</td>
<td>138</td>
</tr>
<tr>
<td>6.5.2 Spacing</td>
<td>140</td>
</tr>
<tr>
<td>6.5.3 Aspect Ratio</td>
<td>140</td>
</tr>
<tr>
<td>6.5.4 Magnet Footprint</td>
<td>142</td>
</tr>
<tr>
<td>6.5.5 Magnet Thickness</td>
<td>143</td>
</tr>
<tr>
<td>6.5.6 Material</td>
<td>144</td>
</tr>
<tr>
<td>6.6 Energy Trends</td>
<td>145</td>
</tr>
<tr>
<td>6.6.1 Methodology</td>
<td>146</td>
</tr>
<tr>
<td>6.6.2 Spacing</td>
<td>147</td>
</tr>
<tr>
<td>6.6.3 Aspect Ratio</td>
<td>147</td>
</tr>
<tr>
<td>6.6.4 Magnet Footprint</td>
<td>149</td>
</tr>
<tr>
<td>6.6.5 Magnet Thickness</td>
<td>149</td>
</tr>
<tr>
<td>6.6.6 Material</td>
<td>151</td>
</tr>
<tr>
<td>6.7 Stability vs. Energy Trends</td>
<td>151</td>
</tr>
<tr>
<td>6.7.1 Spacing</td>
<td>151</td>
</tr>
<tr>
<td>6.7.2 Aspect Ratio</td>
<td>152</td>
</tr>
<tr>
<td>6.7.3 Magnet Footprint</td>
<td>154</td>
</tr>
</tbody>
</table>
FIGURES

2.1 (a) Schematic representation of a device in isolation and definition of coordinate system. (b) SEM of demonstrated anti-ferromagnetically ordered line [36]; (c) MFM of demonstrated anti-ferromagnetically ordered line [36]; (d) SEM of demonstrated single-plane clock wire with nanomagnets [3].

2.2 The role of the clocking field: (a) initial, correct, AF-ordered state; (b) a new input is applied to the line and an external magnetic field biases magnets toward their hard axes (switching may begin at this point depending on clock field strength); (c) additional magnets are put into a logically correct state by their neighbor’s bias and when the applied field is removed they settle to a strong AF-ordered state; (d) Proposed clocking structure: a copper wire clad with a ferromagnetic material (e.g., permalloy) directs flux through magnets to clock a circuit as in (a)-(c).

2.3 (a) Energy minima assuming different biases. Magnets at the end of a clock line may be in an (undesirable) high energy state; (b) Given the average x-bias generated by one magnet and the clock, a y-bias of approximately 23mT is needed to induce an $M_y$ state transition. If a weaker bias is provided (see inset) a device can retain its old state and a stuck-at fault ensues.

3.1 For a $60 \times 90 \times 20 \text{ nm}^3$ permalloy magnet in isolation, note that as the magnitude of the external field ($H_x$) increases, the magnitude of the biasing field (applied in the y-direction) required to flip the magnet decreases.

3.2 The simulation described by a solid line curve without markers illustrates the external field required to place a $60\times90\times30 \text{ nm}^3$ magnet in isolation in a $0^\circ$/metastable state. The line with markers considers the same magnet coupled with a block of magnetic material. A lower external field is required to facilitate a state change.

3.3 As a magnet’s aspect ratio increases, given a similar biasing field, the magnitude of $H_{\text{clock}}$ required to null/switch the magnet increases.

3.4 The magnitude of the y-component of the field produced by a $60 \times 90 \text{ nm}^2$ magnet is measured at two different surfaces.

3.5 (a) Different groups of magnets will be controlled by different metal wires as discussed in Ch. [22]. (b)-(d) Wires can be active at different times to facilitate dataflow in the nanomagnets.
3.6 $M_y$ of the 5th magnet in the line of nanomagnets shown in the insets as a function of $M_x$ of the 6th magnet in the line. Note that the magnets in wire group (i) only switch successfully if the magnetic material at the clock wire boundary is sufficiently nulled. [61]. ........................................... 31

3.7 $M_y$ of two different arrangements of nanomagnets as a function of $H_{clock}$ applied along the magnets’ hard axes. [61]. .......................... 33

3.8 Vertical dataflow Design 1: before (a), after (b); Vertical dataflow Design 2: before (c), after (d). [61]. ........................................... 36

3.9 $M_y$ of the output of two corner turn designs vs. time. [61]. ...................... 37

3.10 Magnitude and direction of field between magnets of crossings vs. time. [61]. ........................................... 39

4.1 (a) AF-ordered line controlled by different clock wires; (b) Improper clocking can lead to mistake states; (c) Field strength 15 nm above conductors in the x- and z-directions when the first two conductors are in a steady state. [25]. ........................................... 43

4.2 A driver magnet to an internal clock wire group may provide less y-directed flux on a neighboring device (dashed line) than the initial input to an NML ensemble (solid line). (Plotted here is the absolute value of the y-directed flux 12 nm away from a $60 \times 90 \times 30$ nm$^3$ supermalloy device across a 60 nm long line.) [25]. ........................................... 45

4.3 (a) For realistic clock distributions, higher external fields are required to facilitate a logically correct state transitions at the second-third clock wire boundary; (b) In an annealed yoke, field lines should not upset magnetic device state. [24]. ........................................... 49

4.4 Magnetic field orthogonal to the direction of current flow along the surface of a 2 micron wide wire (assuming 4 different wire thicknesses). [24]. . . 52

4.5 Multi-plane clocking with a single plane of nanomagnets. Copper is gray, magnetic cladding is black. Dashed vertical line indicates measurement point for overlap of two wires. ........................................... 56

4.6 Multi-plane clocking with two planes of nanomagnets. Red Xs, circles denote direction in which current is applied but not any specific excitation pattern. ........................................... 58

4.7 Field strength as a function of position for the seven initial multi-plane clocking designs. Note that line markers do not denote data points (data was gathered every 1 nm in the x-direction) but rather distinguish different lines. ........................................... 61
4.8 Fields generated when the lower left and center wires are excited (see inset, red X). Aspect ratio of all wires is 1.0. Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.

4.9 Fields generated in NML Level 2. Aspect ratio of all wires is 1.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.

4.10 Absolute value of fields generated in NML Level 2 by excitation patterns in Fig. 4.8 and Fig. 4.9. Aspect ratio of all wires is 1.0.

4.11 Fields generated when the lower left wire is excited (see inset, red X). Aspect ratio of all wires is 1.0. Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.

4.12 Fields generated in NML Level 1 when the lower left and center clock wires (aspect ratio of all wires is 3.0) are excited (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.

4.13 Fields generated in NML Level 2. Aspect ratio of all wires is 3.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.

4.14 Absolute value of fields generated in NML Level 2 by excitation patterns in Fig. 4.12 and Fig. 4.13. Aspect ratio of all wires is 3.0.

4.15 Fields generated in NML Level 2 when the lower left clock wire (aspect ratio of all wires is 3.0) is excited (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
4.16 Fields generated in NML Level 2 for currents of 6 mA in all but the center wire where current is 11 mA. Aspect ratio of all wires is 1.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed. ............................................ 75

4.17 Fields generated in NML Level 2 for currents of 6 mA in all but the center wire where current is 11 mA. Aspect ratio of all wires is 3.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed. ............................................ 76

5.1 Excitation patterns for (a) 4-phase clock, (b) 3-phase clock, and (c) 2-phase clock. \( t_\phi \) is the phase time, i.e., the time duration of each phase. The red (solid) and blue (dashed) curves show the excitation pattern for two adjacent clock wires. The devices patterned on the clock wires represented by the solid red curve will drive the devices on the wires represented by the dashed blue curve. [25]. ......................... 81

5.2 (a) Comparison of forward biases (light curves) and backward biases (dark curves) for trapezoid shaped magnets (dashed curves) and rounded-rectangle (RR) shaped magnets (solid curves). Plotted here is the absolute value of the y-directed flux across a 40 nm line 10 nm away from a given magnet shape; (b) Trapezoids at the boundaries between clock groups to prevent back propagation; (c) Comparison of fields required to place a single trapezoid shaped magnet (dark curve) and rounded rectangle shaped magnet (light curve) into a metastable state. Note the trapezoid has a higher aspect ratio than the rounded-rectangle; (d) Comparison of fields required to place a single trapezoid shaped magnet (dark curve) and rounded rectangle shaped magnet (light curve) into a metastable state. In this case, both magnets have the same aspect ratio. [25]. ......................... 88

5.3 NML majority gate adder schematic. Note that the vertical interconnect, wire crossing (inset at left), and majority gates (M1, M2, M3) have all been simulated and/or experimentally demonstrated. [25]. ......................... 99

5.4 Energy barrier for nanomagnet with biaxial anisotropy for various values of \( K_1 \). Data is captured using OOMMF. [26]. ......................... 106
5.5 Snapshots of micromagnetic simulation for line of magnets with biaxial anisotropy (K1 = 30,000 J/m³). The third clock group and block are omitted for readability. Insets highlight current clock phase for first (solid red curve) and second (dashed blue curve) clock groups. Field strength is 25 mT and the 3-phase clock from Fig. 5.1b is employed. (a) The first clock group is in the “buffer” state (clock field applied), the second is in the “driving” state (no field applied). (b) The first clock group is in the “switching” state (clock field applied), the second is in the “buffer” state (clock field applied). (c) The first clock group is in the “driving” state (no field applied), the second is in the “switching” state (clock field applied). 108

5.6 Bit definition in a pipelined, AF-line for a 3-phase clock. 108

5.7 Comparison of 60 × 90 × 5 nm³ cobalt device with biaxial anisotropy (K1 = 30,000 J/m³) to cobalt device with K1 = 0 J/m³ for increasing hard-axis field strength. +/- markers denote minima on each curve. 110

5.8 3-phase clock excitation pattern for use with biaxial devices controlled by a line clock. 111

5.9 Snapshots of micromagnetic simulation for line of magnets with biaxial anisotropy (K1 = 30,000 J/m³). The third clock group and block are omitted for readability. Insets highlight current clock phase for first (solid red curve) and second (dashed blue curve) clock groups. Field strength is 25 mT and the 3-phase clock from Fig. 5.8 is employed. (a) The first clock group is in the “buffer” state (clock field applied), the second is in the “driving” state (no field applied). (b) The first clock group is in the “switching” state (clock field only applied briefly to ensure unidirectional dataflow), the second is in the “buffer” state (clock field applied). (c) The first clock group is in the “driving” state (no field applied), the second is in the “switching” state (clock field only applied briefly to ensure unidirectional dataflow). 112

5.10 Comparison of 60 × 90 × 5 nm³ cobalt device with biaxial anisotropy (K1 = 15,000 J/m³) to cobalt device with K1 = 0 J/m³ for increasing hard-axis field strength. +/- markers denote minima on each curve. 115

5.11 Snapshots of micromagnetic simulation for line of magnets with biaxial anisotropy (K1 = 15,000 J/m³). The third clock group and block are omitted for readability. Insets highlight current clock phase for first (solid red curve) and second (dashed blue curve) clock groups. Field strength is 25 mT and the 3-phase clock from Fig. 5.1b is employed. (a) The first clock group is in the “buffer” state (clock field applied), the second is in the “driving” state (no field applied). (b) The first clock group is in the “switching” state (clock field applied), the second is in the “buffer” state (clock field applied). (c) The first clock group is in the “driving” state (no field applied), the second is in the “switching” state (clock field applied). 116
6.1 Conflicts occur if (a) an odd number of magnets per clock group or (b) an even number of magnets per clock group is used. [26].

6.2 Snapshots of random walk in $40 \times 60 \times 20 \text{nm}^3$ line showing (i) the original, conflict state (ii) the migration through the $0^\circ$ state resulting in (iii) the conflict walking to an adjacent device. [26].

6.3 Latch methods: (i) AML (ii) SML (iii) TML.

6.4 Predicted EB ($40 \times 60 \times 20 \text{nm}^3$ device, 8 nm spacing). [26].

6.5 Predicted EB ($40 \times 80 \times 20 \text{nm}^3$ AML, 8 nm spacing). [26].

6.6 Predicted EB ($40 \times 80 \times 20 \text{nm}^3$ SML, 8 nm spacing). [26].

6.7 3-phase clock for use with SML and TML. A high field is used in the first phase to place the latch in a metastable state. The field is then lowered (e.g., between 0-5 mT in this work) to clock the remaining magnets.

6.8 Stability as a function of spacing for the TML.

6.9 Stability as a function of aspect ratio for the TML.

6.10 Stability as a function of magnet footprint for the TML.

6.11 Stability as a function of magnet thickness for the TML.

6.12 Energy (estimated by mT$^2$) as a function of spacing for the TML. Note that at an 8 nm spacing the lines of cobalt magnets with $K_1 = 30,000 \text{J/m}^3$ could not be clocked.

6.13 Energy (estimated by mT$^2$) as a function of AR for the TML.

6.14 Energy (estimated by mT$^2$) as a function of magnet footprint for the TML.

6.15 Energy (estimated by mT$^2$) as a function of thickness for the TML.

6.16 Stability (kT) per energy (estimated by mT$^2$) as a function of spacing for the TML. Recall that for an 8 nm spacing, the lines of cobalt magnets with $K_1 = 30,000 \text{J/m}^3$ could not be clocked.

6.17 Stability (kT) per energy (estimated by mT$^2$) as a function of aspect ratio for the TML.

6.18 Stability (kT) per energy (estimated by mT$^2$) as a function of magnet footprint for the TML.

6.19 Stability (kT) per energy (estimated by mT$^2$) as a function of thickness for the TML.

7.1 (a) Serial implementation of stochastic multiplier via a single 2-input AND gate. (b) Parallel implementation of stochastic multiplier via 10 2-input AND gates. (c) Serial implementation of stochastic scaled addition via multiplexer.

7.2 Block diagram of circuit to implement a degree-3 Bernstein polynomial as described in [75].
7.3 Comparison of iNML (left) to PML (right). [63]. ................................. 176
7.4 Energy and delay for deterministic (CMOS) and stochastic (CMOS, PML) multipliers. Only low-power CMOS is considered. ................................. 184
7.5 Area for deterministic (CMOS) and stochastic (CMOS, PML) multipliers. Only low-power CMOS is considered. ................................. 184
7.6 Area-energy-delay product for deterministic (CMOS) and stochastic (CMOS, PML) multipliers. Only low-power CMOS is considered, and only the lowest AEDP of the various voltages (0.3 to 0.9 V) is displayed for each PTM data point (for clarity). ................................. 185
7.7 PML circuit to implement a degree-3 Bernstein polynomial per the method in [75]. ................................. 188
TABLES

3.1 $M_{y\text{--avg}}$ AND $H_{y\text{--Bias}}$ FOR MAGNETS OF DIFFERENT ASPECT RATIOS. ................................................................. 26

4.1 POWER PER WIRE ASSUMING 25 PARALLEL LINES ATOP EACH CLOCK WIRE. ................................................................. 49

4.2 SUMMARY OF DESIGNS USED TO STUDY THE EFFECTS OF CHANGING HORIZONTAL AND VERTICAL SPACING BETWEEN TWO MULTIPLE PLANE CLOCK WIRES AS IN FIG. 4.5 .................................................. 60

4.3 SUMMARY OF FIELD DISTRIBUTIONS FOR CLOCK WIRES WITH ARS OF 1.0 AND 3.0. ................................................................. 71

5.1 TRENDS ASSOCIATED WITH MAGNET SPACING AND SIZE WITH RESPECT TO THEIR EFFECTS ON THE MINIMUM REQUIRED FIELD STRENGTH AND PHASE TIME. ................................................................. 92

5.2 TRENDS ASSOCIATED WITH SCALING THE WIDTH OF THE YOKE BETWEEN COPPER CLOCK WIRES. ................................................................. 96

5.3 PERFORMANCE PROJECTIONS FOR 32-BIT SUBTHRESHOLD CMOS ADD. ................................................................................. 101

5.4 PERFORMANCE PROJECTIONS FOR 32-BIT NML ADD. ............. 103

7.1 COMPARISON OF AREA DELAY PRODUCT (UNITLESS) FOR IMPLEMENTATIONS OF A DEGREE-3 BERNSTEIN POLYNOMIAL. . . . . . 190
ACKNOWLEDGMENTS

First, I would like to acknowledge and thank my advisor Dr. Michael Niemier, as well as Dr. Sharon Hu, for their guidance and support throughout my time at Notre Dame. Thank you also to Dr. Gyorgy Csaba and Dr. Wolfgang Porod for their support and for serving on my committee.

I would also like to thank the members of our research group, especially Dr. Steve Kurtz, with whom I worked closely. I am also grateful to the Semiconductor Research Corporation and Nanoelectronics Research Initiative, to the National Science Foundation, and to the University of Notre Dame for supporting my research.

Finally, thank you to all my family, especially my wife Ashley, for their continual love and encouragement.
CHAPTER 1

INTRODUCTION

1.1 Overview

This work considers nanomagnet logic (NML), an emerging device technology that is the subject of many recent research efforts, including the Semiconductor Research Corporation (SRC) Nanoelectronics Research Initiative (NRI). As fundamental scaling limits for the complementary metal-oxide-semiconductor (CMOS) field-effect transistor (FET) are on the horizon, the goal of the NRI is to investigate emerging technologies to replace or augment CMOS. In NML, nanoscale islands of magnetic material encode binary information via magnetic polarization. Computing in NML is accomplished through nearest-neighbor interactions between these nanomagnets. NML devices can be envisioned in a quantum dot cellular automata (QCA) architecture [45], and thus in the past was called magnetic QCA (MQCA). (In accordance with the [37], I use the term NML, even in text that is incorporated from past publications.)

There are many potential benefits NML offers over CMOS: devices above the superparamagnetic limit are inherently non-volatile (potentially enabling instant-on operation and inherent checkpointing), switching device state is a low-power operation, and magnetic devices are intrinsically radiation hard. This is not to say that there are not possible drawbacks to NML. For example, computation speeds in NML are limited by the relatively slow switching speed of devices, and thus it cannot effectively serve in the same roles as high performance CMOS. Therefore, device characteristics push NML toward certain application spaces. As just one example, space-based information processing hardware
would benefit from non-volatility, radiation hardness, and low-power operation, but would not necessarily be limited by relatively slow switching speeds.

Alternative forms of QCA have also been proposed, but studies suggest NML may be the most promising of any implementation of the QCA architecture due to limitations in the other approaches. For example, metal-dot QCA structures have been physically implemented (devices [70], gates [5], fanout [87], etc.), and molecular QCA has been proposed [38, 74]. These methods suffer from a variety of problems that greatly limit their feasibility. For molecular QCA [74], there is currently no device that can be deterministically placed, and small variations, stray charges, etc. severely limit functional correctness [64]. Similarly, the low operating temperatures for electrostatic QCA [5] limit application spaces.

Given the above, the main goal of my work is to elevate the understanding of NML from the device and circuit levels to allow for the study of applications and architectures where NML can offer performance improvements compared to CMOS equivalents. I now briefly outline each of my contributions in this regard, and describe the contents of this document.

1.2 My Contributions

I now describe the main contributions I have made to the state of the art for NML circuits, applications, and architectures as described in this document.

1.2.1 NML Circuits

One of the main goals of NML research, and my work specifically, is to show that there can be performance gains at the application level, i.e., when compared to functional equivalents in state of the art CMOS. Prior to the onset of my work, the majority of NML research was at the device level (i.e., simple logic gates, etc. in isolation both in simulation and experiment). Additional analysis and design for NML circuits are required in order to
bridge the divide between device and application level study. In Ch. 3, I present methods that further the understanding of NML devices, and facilitate the transition from the device to circuit level, including:

- a method to represent the magnetization state of NML devices using non-traditional MH curves to enable studies of device-to-device interaction (necessary to design more complex structures).
- a study of how the shape and size of a device affect its ability to drive and to be driven by other devices – especially important given that all interaction between devices is via nearest-neighbor interaction.
- a study of how information is transferred across boundaries between clock wires (i.e., as the proposed method to re-evaluate the magnets employs clad, current-carrying copper wires that each control many NML devices), and how to eliminate unwanted feedback in clocked circuits.
- a simulation-based case study demonstrating how the aforementioned methods can be used to design a variety of NML circuit elements necessary for higher-level constructs (including inverter chains, wire crossings, and corner turns).

1.2.2 Modeling Realistic Clock Wires

As my work focuses on the study of physically-implementable clock structures, it is important to understand the power consumption of this clocking method. Despite the fact that NML devices dissipate low power while switching, and thus could be competitive with CMOS, we must also account for the energy requirements of the clocking mechanism required to facilitate re-evaluation of NML ensembles. The proposed clock structure considered in my work could dwarf NML devices in terms of power consumption. Therefore, in Ch. 4 I describe:

- methods to quantify the power and energy requirements of this proposed clock in order to compare NML-based logic to CMOS functional equivalents.
- a simulation-based method to model the realistic clock structures.
- how field distributions generated by realistic clock wires impact logical correctness in NML circuits.
- how defects and variation in the fabricated clock wires can affect circuit operation and power requirements.
• a study of the effects of interference from clock structures on the fields generated by their neighbors. This study is performed in the context of single-plane clock wires, and multi-plane clock wires (where we could potentially implement logic in multiple planes of NML).

1.2.3 Uni-directional Dataflow, Energy Analysis, and Comparison to CMOS

In addition to these first steps there are other critical considerations for NML circuit design. In Ch. [5] I present:

• a method for addressing one of the five tenets of digital logic (proposed in [86]), the ability to guarantee unidirectional dataflow in large clocked systems. In any circuit, data must flow from inputs to outputs, but not from outputs to inputs. As such, a concern with proposed implementations of QCA is how to guarantee unidirectional dataflow; this is also a key challenge for NML which I addressed for the first time (in the context of realistic clock structures). Namely, I show how one can employ multi-phase clock waveforms and device shape engineering to ensure unidirectional dataflow in NML circuits.

• a continuation of the study from Ch. [4] to consider clocking theory, energy, and delay in more detail, which enables me to make performance projections to quantify energy and delay for an NML ripple carry adder. Notably, I demonstrate that NML can achieve as much as an approximately 100X improvement in terms of the energy-delay product over equivalent CMOS circuits.

• consideration of materials-based solutions to achieve even lower energy operation. With respect to the material of the magnets, I compare the performance and energy of two distinct clocking approaches (switching in the presence vs. the absence of an applied field) in the context of multi-phase clocks.

1.2.4 NML Shift Registers and Non-volatility

The majority of my study has focused on NML inverter chains, which can be used for interconnect (a primary emphasis has been on local interconnect, as NML devices switch relatively slowly), but also for other tasks. A non-volatile NML inverter chain can be used in a shift-register memory architecture. The shift register is well-studied, and arguably the most simple and robust NML circuit imaginable. Storage in computing does not necessarily need to be non-volatile, unless it stores data that needs to be kept across
power cycles. Traditional computing systems use volatile memory at all but the highest levels, saving critical data to non-volatile disk-based storage to maintain it when powered off. However, many applications could benefit from non-volatile storage at lower levels of memory, e.g., main memory, cache, or even the register level. A variety of these application spaces can be envisioned for an NML shift-register memory architecture.

In Ch. 6 I consider an NML-based shift register. One hurdle that needs to be overcome to enable a shift-register memory architecture is ensuring non-volatility in the lines. A typical NML device may be non-volatile in isolation, but ensemble interactions due to nearest-neighbor coupling can lower a device’s energy barrier such that its state is susceptible to thermal noise or stray fields. Therefore, in Ch. 6 I describe:

- how a destructive random walk can arise even in defect- and error-free, pipelined circuits (e.g., shift registers) when power is removed from the system.
- a review of existing methods to ensure stability of NML devices in ensembles (based on work in [43]), and an approach for how to leverage these methods to provide non-volatility in NML shift registers.
- a design strategy to identify designs that can offer high stability and low energy, novel clock waveforms to enable these designs, and a design space study to identify specific designs that offer high stability and low energy.
- a scheme in which NML shift registers could be used: a general checkpointing architecture. Non-volatile shift registers can store computational state which will be available to restart computation in the case of power failure. Similarly, a soft error could require computation to be restarted from the stored, stable state. I therefore also discuss the feasibility of a checkpointing application, including preliminary performance comparisons to existing methods and paths of future study.

1.2.5 Stochastic Computing

Finally, in Ch. 7 I perform another application-level study enabled by my initial work. In stochastic computing, computation is performed on data encoded in streams of random bits [33, 32]. Stochastic computing offers high performance for certain applications, e.g., image processing, and is highly tolerant of errors [46, 52, 75, 77]. Typically, the random bit streams are organized serially, but streams could also be organized in parallel. An NML
implementation of a stochastic circuit lends itself to an efficient parallel implementation as many bits could be controlled by a single line clock, for example. Therefore, in Ch. 7 I describe my work to:

- quantify performance (in terms of area, energy, and delay) for an NML implementation of a stochastic multiplier, and compare to deterministic and stochastic CMOS equivalents.
- design an NML implementation of a circuit described in [75] that can implement a number of stochastic logic functions, and compare this design to the CMOS equivalents in [75] in terms of area, delay, and the area-delay product.

1.3 Organization of Document

To summarize, this document is organized as follows. In Ch. 2 I discuss NML state of the art, and introduce the five tenets for digital logic proposed by [86]. I also discuss a clock structure for NML that has been fabricated and experimentally validated, and explain how bidirectional dataflow can arise in NML circuits driven by these structures. In Ch. 3 I present the studies to bridge the gap between the device and circuit levels. Next, in Ch. 4 I present my work to quantify the impact of realistic clock wires on logical correctness and power consumption, and to quantify the effects of defects, variation and field interference from these clock wires on NML circuit operation. In Ch. 5 I present theory to quantify energy and delay of NML circuits, show how multi-phase clocks can ensure uni-directional dataflow in NML circuits, and present performance projections to compare to a CMOS ripple carry adder (including a discussion of how to achieve lower-energy operation via new magnetic materials, etc.). Next, in Ch. 6 I present my work relating to NML shift registers, including how to enable low energy, high stability designs, and a discussion of a checkpointing application. Finally, in Ch. 7 I describe my work to quantify performance of NML for use in a stochastic computing architecture. Ch. 8 concludes.
CHAPTER 2

BACKGROUND

In this section I review fundamental background for NML. I discuss possible clocking mechanisms and highlight the clocking method that is central to my work.

2.1 Fundamental Constructs

Binary state can be represented by the polarization of single-domain nanomagnets (typically composed of permalloy, an alloy of Ni and Fe, or supermalloy, an alloy of Ni, Fe, and Mo). A nanomagnet with an aspect ratio above 1.0 (i.e., taller than it is wide) will have a preferred axis of magnetization in the y-direction (see Fig. 2.1a) due to shape anisotropy. This creates two stable states for the device in the y-direction (i.e., “up” and “down”) one of which can represent a logic ‘0’, and the other a logic ‘1’. Horizontally adjacent nanomagnets couple with each other anti-ferromagnetically (AF) – see Fig. 2.1b-c – while vertically adjacent nanomagnets couple ferromagnetically (F). This coupling can be used to propagate data and perform logic operations. Interconnect is formed by lines of either horizontally or vertically coupled nanomagnets – both AF- and F-ordered lines have been demonstrated experimentally. The basic logic gate in NML is the majority voting gate in which output is determined by the majority of three inputs. The majority gate can be modified to perform a logic AND or OR operation by fixing one input to a logic ‘0’ or ‘1’, respectively. Early fabrication experiments of devices, wires, and majority gates at room temperature was first presented in [36].

1Magnets need not be oblong if other forms of anisotropy are leveraged, e.g., magnetocrystalline anisotropy, which is based on atomic structure.

7
2.2 Clocking

The energy difference between the two stable magnetic polarizations of NML devices can be large, necessitating an external stimulus to allow for transition between states in accordance with inputs. The external stimulus can be provided in the form of a magnetic field applied along the hard axes of the devices. Logical re-evaluation follows these steps: the ensemble starts in an AF-ordered, correct state from previous computation (Fig. 2.2a); next, the input is set and the external field is applied, lowering the energy barrier between stable states and allowing the signal to begin to propagate (Fig. 2.2b); finally, the external field is removed and the devices completely settle to the new state as set by the input (Fig. 2.2c).

The external clocking stimulus can be created in a number of ways. The primary method considered in this work was first proposed in [59], in which a current-carrying copper wire is wrapped in a ferromagnetic material such as permalloy (see Fig. 2.2d). Current through the wire creates a magnetic field above the surface of the wire where NML...
Figure 2.2. The role of the clocking field: (a) initial, correct, AF-ordered state; (b) a new input is applied to the line and an external magnetic field biases magnets toward their hard axes (switching may begin at this point depending on clock field strength); (c) additional magnets are put into a logically correct state by their neighbor’s bias and when the applied field is removed they settle to a strong AF-ordered state; (d) Proposed clocking structure: a copper wire clad with a ferromagnetic material (e.g., permalloy) directs flux through magnets to clock a circuit as in (a)-(c). [26].
devices would be placed. The ferromagnetic cladding acts to direct increased magnetic field over the NML devices. The maximum number of clocked devices that can be controlled by a single clock wire has not been well-studied, but simulation suggests it could be on the order of tens of nanomagnets (see Ch. 5.3). As such, multiple clock wires will need to be placed adjacent to each other, as in Fig. 2.2d. A primary goal of this work is to understand the impact of multiple, adjacent clock structures on NML circuits in terms of circuit robustness, power, delay, and energy. This study is presented in Chs. 4 - 5.

The current-carrying wire approach has been well-studied for use in field magnetoresistive random access memory (MRAM), in which the state of magnetic storage elements are set via fields from these wires (as opposed to using the STT effect, which has recently superseded field-based switching). For NML, clock structures with nanomagnets patterned on top have been created in recent fabrication efforts, and the fields from the wires can switch NML devices [3] as well as NML lines and gates [2]. Note that in [3] the devices were patterned in isolation, and oriented so switching occurred along their easy axes. In [2], the inputs to the lines and gates were similarly oriented and also necessitated very high field strengths. These field strengths and currents thus are not indicative of the fields to switch the NML ensembles studied in this work (to be discussed).

Before continuing, it is worth noting that other approaches to clocking have also been proposed. [8] proposes on-chip solenoids for clocking individual devices, requiring solenoids with 1 nm wide coils. There is currently no viable way to fabricate devices with such small features. NML devices could also be controlled by multi-ferroic materials [78, 16, 31]. For example, materials that exhibit ferromagnetism and ferroelectricity as in [16], or materials that exhibit ferromagnetism and ferroelasticity as in [31], could be individually controlled with small voltage pulses, resulting in low energy operation. However, these approaches require that every NML device be individually contacted, resulting in increased fabrication complexity over the current-carrying clad wire approach. Furthermore, the line clock approach has been shown to switch devices [3], and lines and gates
2. As such, this work primarily focuses on the current-carrying wire approach as it is well-studied both in simulation and experiment. Furthermore, as will be demonstrated in Ch. 5, the energy requirements of these clock structures should be low enough to allow significant performance gains compared to equivalent CMOS circuits.

2.3 I/O

Converting the magnetic state of a device to an electric signal, and vice versa, is a subject of continuing study. The proposed method for input (i.e., converting an electric signal to a magnetic state) is to use the spin-transfer torque (STT) effect to set the state of a device. The principle of the STT effect is that if spin-polarized electrons pass through a magnetic material they will exert a torque that can be used to set the state of the device. This method is similar to that used in STT-based MRAM, and is well-studied. This method was experimentally demonstrated in [55], which considered circuits where devices are composed of magnetic tunnel junctions (MTJ) (rather than a single plane of magnetic material). Another proposed input method is to fabricate a small wire over the input magnet which will produce a nominal magnetic biasing field when current is directed through it [80]. If the energy barrier between states is sufficiently low, the biasing field can tip the input to the desired direction.

The proposed output method (i.e., converting magnetic device state to an electric signal) for NML is again modeled after a scheme used in MRAM. The tunneling magnetoresistive (TMR) effect states that if two magnetic layers are separated by a small (a few nanometers) insulating layer, and a current is passed through them, electrons are able to tunnel from one magnetic layer to the other. If the two ferromagnetic layers are ferromagnetically aligned, they produce a higher resistance than if antiferromagnetically (AF) aligned. The stack of contacts, magnetic layers, and insulator is called an MTJ. The TMR effect can be used in NML to read magnetic state – the state of the free layer of an MTJ is set by fringing fields from neighboring nanomagnets, and its orientation with respect to
the non-free layer results in a resistance that can be measured to determine state. Recent work on MTJs for NML output is presented in [53, 54], and is a subject of ongoing study both in simulation and experiment.

2.4 Tenets of Digital Logic

As discussed in Ch. 1, [86] presents five tenets a digital logic device should satisfy if it is to be used in a digital system. The five tenets are that the device should: 1) be able to implement a functionally complete logic set; 2) provide power amplification; 3) exhibit concatenability; 4) have non-linear response characteristics; 5) have data flow unidirectionally.

In this section, how NML satisfies tenets 1-4 is briefly explained. For tenet 5 – which is addressed for the first time in this work – how unwanted feedback arises in NML circuits is described. Ch. 5 describes how circuits can be designed to avoid this feedback in the context of realistic, local clock structures.

2.4.1 Functionally Complete Logic Set

A digital logic system cannot exist without a functionally complete logic set, i.e., a set of logic operations that can be used in conjunction to implement any higher-level logic operation. Boolean inversion (\texttt{NOT}) and either the \texttt{AND} or \texttt{OR} operation implement a functionally complete logic set. As described in Ch. 2.1, all three of these logic operations have been implemented in NML, and validated through simulation and experiment.

Reduced footprint logic gates have also been proposed for NML. The \texttt{AND} and \texttt{OR} gates based on the modified majority gate require three input magnets, one decision magnet, and an inverting output magnet (since NML actually implements an inverting majority function). Recent work has proposed shape-based logic gates that could perform \texttt{AND/OR} logic with a reduced footprint [62, 84]. The shape-based gate requires one less input, and the resulting decrease in area and critical path length (and thus delay) could impact system-
level performance significantly (e.g., the critical path in an $\mathrm{XOR}$ decreases almost 30% from 14 to 10 magnets in $\text{[62]}$). The shape-based gates operate by altering the energy landscape of the devices, such that they prefer one direction of $y$-magnetization over the other. For example, an oval-shaped NML device with a slanted edge on the upper righthand corner gives it a preferred magnetization in the “up” direction. Thus, when one of the two inputs is “up”, the output is also “up” due to ferromagnetic coupling and preference imparted by the slant. Only when both inputs are oriented “down” will the slanted magnet transition to a “down” state. If “up” corresponds with a logic ‘1’ and “down” with a logic ‘0’, this is an $\mathrm{OR}$ function. Similarly, an $\mathrm{AND}$ function can be produced by placing the slant on the lower left.

2.4.2 Power Amplification

This tenet states that a circuit must exhibit power amplification, or gain, from inputs to outputs. That is, a circuit must be able to effectively drive multiple outputs, and the signal must be maintained when transferred over long distances. As proof-of-concept, a fanout of 1:3 was experimentally demonstrated in $\text{[83]}$. Amplification can be provided by the clocking field. For example, consider re-evaluation of a line of magnets by an external clocking field (as in Fig. 2.2a-c). As the clock is applied, the energy barrier of the magnets is lowered such that the first magnet in the line is biased by the state of the input, followed by the second magnet in the line being biased by the first magnet, etc. As the clock field is removed, the magnets relax to a strong “up” or “down” state and they become strong drivers for subsequent devices.

2.4.3 Non-linear Response

A non-linear response is required in order to achieve a high signal-to-noise ratio in large circuits. The response characteristics of NML devices are inherently non-linear, as verified by simulation and experiment. Experimental validation is presented in $\text{[82]}$. In
Ch. 3.1.2 – in which methodologies for analyzing NML circuit elements are presented – simulations of an NML device in isolation shows non-linear state transition in the presence of an applied field (i.e., the clock) and a small biasing field (i.e., to mimic the influence of a neighboring device).

2.4.4 Concatenability

Signals from multiple NML devices must be concatenable, meaning that the output of a device must be compatible with the input. In NML, the fringing fields from one device are able to set the state of other devices whose energy barriers are modulated by the clock. As discussed previously, experimental results show correct operation of interconnect, fan-out, majority gates, etc.

2.4.5 Uni-directional Dataflow

This tenet states that data must flow only from inputs to outputs, i.e., there must be no unwanted feedback. This tenet is considered in the context of NML – specifically, in the context of physically-realizable clock structures – for the first time in this work. Unlike electrostatic implementations of QCA – where the clock dictates the direction of data propagation in an ensemble of QCA devices [34] – in a magnetic implementation, the direction in which an external field is applied along the magnets’ hard axes does not necessarily matter. For example, referring to Fig. 2.2b, the second and third magnets in the AF line could also be biased from right-to-left (i.e., ←) and the line would still settle into the AF-ordered state illustrated in Fig. 2.2c. This makes it much easier for feedback to occur. Data can flow in multiple directions given the same, uni-directional, external magnetic field. In order to satisfy this final tenet for digital logic, it must be ensured that data propagates in the correct direction to prevent mistake states caused by unwanted feedback.

Since NML ensembles need to be controlled by multiple clock lines, it is imperative
that each hard-axis biased magnet in the ensemble (e.g., the line in Fig. 2.2a-c) see only one strong driver. When considering an “internal” magnet in an AF-ordered line, the clock field and flux from neighboring magnets keep individual devices biased along their hard axes until they are switched by the appropriate neighbor. As an example, if two $60 \times 90 \times 30$ nm$^3$ supermalloy magnets are spaced 25 nm apart, an external field of just $\sim 20$ mT (in addition to the fields provided by the magnets themselves) would be needed to ensure that the energy minimum of a device between them remains at 0 degrees (see ‘+’ magnet in Fig. 2.3a) and closer magnet spacings would further reduce this field. While in this state, a y-bias from a neighboring device can induce a state transition (see Fig. 2.3a where the energy minimum for the ‘*’ magnet is shifted to $\sim 40$ degrees due to the bias from the neighbor).

A more difficult case is ensuring the last magnet in a clocked group only sees one strong driver. Even if the last magnet in an AF-ordered line is already hard-axis biased, it will see only x-directed flux from the clock and one neighbor. As seen from the top dashed curve in Fig. 2.3a, a magnet in the “#” footprint, would be in a high-energy state, and fabrication variations, temperature effects, etc. could cause the device to flip before it is biased into a correct state by the correct neighboring device. The simulation results in Fig. 2.3b also suggest that flux from one magnet plus the clock could be insufficient to induce a state transition. If the neighbor magnet cannot produce a sufficient bias – e.g., due to a “conflicting” bias from the first, unmanaged device in the next group – the last magnet in the clocked group will retain its old state and a “stuck at” fault may ensue.

Note that for similar reasons, increasing the magnitude of the applied hard axis field will not necessarily improve signal propagation in an NML circuit – as it can make each magnet a weaker driver of its neighbor by increasing said magnet’s x-component of magnetization. Finding ways to avoid both forms of back propagation in NML ensembles – with clock structures that can be implemented on-chip – is a driving goal of this work. In Ch. 5.2, two methods are presented – multi-phase clocks, and shape engineering – that
Figure 2.3. (a) Energy minima assuming different biases. Magnets at the end of a clock line may be in an (undesirable) high energy state; (b) Given the average x-bias generated by one magnet and the clock, a y-bias of approximately 23mT is needed to induce an \( M_y \) state transition. If a weaker bias is provided (see inset) a device can retain its old state and a stuck-at fault ensues. [25].

ensure data propagates only in the desired direction.
CHAPTER 3

FROM NML DEVICES TO CIRCUITS

In this chapter I present generic methods of analysis for NML circuits that is applicable to all NML study, and used throughout my work (Ch. 3.1). This includes: a discussion of two ways to represent the magnetization state of devices (Chs. 3.1.1, 3.1.2) to enable study of device-to-device interaction (necessary to design more complex structures); a study of how the shape and size of a device affects its ability to drive and to be driven (Ch. 3.2); a preliminary study of how information is transferred across clock wire boundaries, and how to eliminate unwanted feedback in clocked circuits (Ch. 3.3). Next, I use these methods to perform a simulation-based feasibility study of specific NML circuits (Chs. 3.4, 3.5).

3.1 Analyzing Magnetic Circuit Elements

Here, a methodology is presented for analyzing magnetic circuit elements that can be used to help generate initial designs, and to evaluate circuit-level functionality and robustness. The OOMMF simulation suite [27] is leveraged. OOMMF, created by NIST, solves the Landau-Lifshitz equation. OOMMF is widely used and there is excellent correlation between simulation and experimental results (see [85, 57]).

3.1.1 Magnetization State - Traditional MH Curves

To switch the magnetization state of an NML device, one must apply a strong bias in the y-direction opposite to its initial magnetization. The minimum field along the y-axis required to perform this switching can be approximated as $H_{bias} = (N_x - N_y)M_s$, where $N_x$
and \( N_y \) are the demagnetizing factors along the hard and easy axes respectively (see Ch. 3.2 for more detail), and \( M_s \) is the saturation magnetization [68].

As an example, an external field of approximately 130,000 A/m\(^1\) would need to be applied along the easy axis of a \( 60 \times 90 \times 20 \) nm\(^3\) permalloy magnet (with rounded rectangle shape, at 0 K to abstract away from temperature effects) in order to change its magnetization state. That said, in an NML-based circuit, the biasing field required to switch the device needs to come from a neighboring input magnet. However, an equivalently sized device placed 20 nm away from the magnet to be switched would realistically only generate an \( H_{\text{bias}} \) of approximately 10,000 A/m when the y-component of magnetization in the next magnet’s “footprint” is averaged. The disparity between the biasing field required and the biasing field a magnet can produce can be addressed by introducing an applied field \( (H_x) \) along a magnet’s hard axis (i.e., its short axis) that helps to facilitate the switching process (see Ch. 2.2). From Fig. 3.1 it is clear that as \( H_x \) increases, the magnitude of the biasing field required to switch this magnet is reduced considerably.

3.1.2 Magnetization State - NML Representation

From this point forward, the data captured by Fig. 3.1 will be presented via a non-traditional hysteresis curve. Namely, the y-component of magnetization \( (M_y) \) as a function of the applied field \( (H_x) \) will be considered – as this analysis inherently includes information about logical correctness/binary state \( (M_y) \) as well as the local clocking field \( (H_x) – \) hereafter referred to as \( H_{\text{clock}} \). Before discussing specific designs, this analysis is used to show how to leverage magnetic material to change the polarization of neighboring nanomagnets more easily. This facilitates tasks such as dataflow in a line of nanomagnets that is orthogonal to the direction of the clocking field.

More specifically, the results of two simulations performed with the OOMMF sim-
Figure 3.1. For a $60 \times 90 \times 20$ nm$^3$ permalloy magnet in isolation, note that as the magnitude of the external field ($H_x$) increases, the magnitude of the biasing field (applied in the y-direction) required to flip the magnet decreases. \cite{61}. 

\[ \text{Magnetization (M}_y\text{)} \]
ulation suite are presented (see Fig. 3.2). All simulations consider a $60 \times 90 \times 30$ nm$^3$ supermalloy magnet (the material used in [17]) initially oriented down ($\downarrow$) $H_{\text{clock}}$ is applied in the positive x-direction (i.e., left-to-right along the magnet’s hard axis). A constant 15,000 A/m biasing field is applied along the magnet’s easy axis which “tips” the magnet to the opposite polarization.

First, a single magnet in isolation is considered (see “iso, $H_{\text{clock}}$” in Fig. 3.2). The magnet is initially oriented down and its y-component of magnetization is strongly negative ($M_y = M_s = -800,000$ A/m). An external field ($H_{\text{clock}}$) is then applied along the hard axis of this magnet. As $H_{\text{clock}}$ increases, the magnitude of the y-component of magnetization decreases – eventually reaching 0 when $H_{\text{clock}}$ is approximately 50,000 A/m.

Note that in the “iso, $H_{\text{clock}}$” curve in Fig. 3.2, the magnitude of $H_{\text{clock}}$ continues to increase even after $M_y$ changes signs (indicating a $\downarrow$ to $\uparrow$ transition). The net result is that $M_y$ is suppressed and actually begins to decrease (see the “tail” at the top-right of the “iso, $H_{\text{clock}}$” curve). Only when $H_{\text{clock}}$ is eventually relaxed does $M_y$ again saturate (this time at positive 800,000 A/m). In the second half of this simulation, the process discussed above is repeated, but this time the sign of $H_{\text{clock}}$ is changed to simulate an external field applied in the opposite direction. Note that the second half of “iso, $H_{\text{clock}}$” in Fig. 3.2 is essentially symmetrical to the first half.

These results raise two important considerations for NML circuit design. First, the applied field can be too strong. In the simulation discussed above, the biasing field is externally generated, but in an NML-based circuit it would come from the nanomagnets in the circuit. While an external nulling field (i.e., one that moves a magnet’s state into a $0^\circ$, metastable state) can assist with switching and the evaluation of new inputs, it can also impede dataflow by suppressing the ability of a magnet to switch its neighbor (as the

---

2As in [21], the following are assumed: a saturation magnetization of $8.0 \times 10^5$ A/m, an exchange stiffness constant of $1.05 \times 10^{-11}$ J/m, an anisotropy constant $K_1$ of $3 \text{ J/m}^3$, and the default damping constant of 0.5. Each simulation stage (when the magnitude or direction of the applied field changes) was considered complete when the maximum $|d\mathbf{m}/dt|$ dropped below a preset number of degrees per nanosecond.
Figure 3.2. The simulation described by a solid line curve without markers illustrates the external field required to place a 60x90x30 nm$^3$ magnet in isolation in a 0°/metastable state. The line with markers considers the same magnet coupled with a block of magnetic material. A lower external field is required to facilitate a state change. [61].
biasing field generated by the magnet will be of lower magnitude). Second, the symmetry of the “iso, $H_{\text{clock}}$” curve suggests that the direction of $H_{\text{clock}}$ should not affect the direction of dataflow. The state that a magnet switches to is determined solely by the biasing field of a neighboring magnet. Thus, a unidirectional clock could facilitate bi-directional dataflow. Feedback is therefore possible within a single clock wire group (depending on the number of clock phases used), and will need to be well controlled – this is considered in Ch. 5.1.1.

While the main focus of this work is not to optimize the energy efficiency of a given design, one cannot neglect the fact that the “iso, $H_{\text{clock}}$” simulation seemingly indicates that an external field of at least approximately 50,000 A/m is required to null/switch a magnet (irrespective of the biasing field). If a field of this magnitude were truly necessary, \[59\] suggests that a current density of over $10^7$ A/cm$^2$ would be required in the clock wires. This would not only dramatically increase the overall system energy, but heating and electromigration effects would be a practical concern.

That said, the magnets and magnetic material that make up a circuit can actually assist in the switching process which will allow one to decrease the magnitude of $H_{\text{clock}}$. To illustrate, again considered is a $60 \times 90 \times 30$ nm$^3$ magnet, but this time a block of magnetic material is added 25 nm to the right of it. The block of magnetic material has an aspect ratio that is less than 1 and its easy axis is parallel to the applied field (as such, its magnetic state should not change). The first half of the simulation discussed earlier in this section (for “iso, $H_{\text{clock}}$”) was repeated with the same 15,000 A/m bias. Looking at the “coupled, $H_{\text{clock}}$” curve, one can see that the magnet switches when $H_{\text{clock}}$ is approximately 26,000 A/m – a field half that required to null/switch a magnet in isolation.

When the block is placed in close vicinity to the nanomagnet, the block concentrates flux towards itself. In other words, as the block is already magnetized along the positive x-direction (due to its shape anisotropy) it increases the magnetic flux density in the x-direction in the surrounding area. This increase in flux density assists the switching process as it creates a strong local magnetic field in the x-direction. In fact, one can consider the
block used in the “coupled, $H_{clock}$” simulation in isolation and calculate the average field generated by the block where the magnet eventually resides. Simulation data indicates that the block provides a local field where $H_x$ averages to be approximately 37,000 A/m. Thus, introduction of the block helps reduce the minimum applied field requirement, which ultimately reduces the external energy needed to perform the switching.

The magnetic field produced by the block magnet and $H_{clock}$ is referred to as $H_{local}$ (thus $H_{local} = H_{clock} + H_{block}$). The magnitude of $H_{local}$ clearly will depend on the block size, aspect ratio, shape, material and distance from the nanomagnet. $H_{local}$ is an important parameter for NML circuit design and should be carefully tuned using all available parameters. Although one might be tempted to think that it is desirable to maximize $H_{local}$ so that $H_{clock}$ can be minimized, as seen earlier, if the net field applied to a magnet is too strong, it can keep a magnet magnetized in the x-direction and impede the biasing field that will ultimately switch a neighboring device.

### 3.2 Shape, Size, and Circuits

While many of the circuit designs discussed here consider sizes and shapes consistent with the work in [59], how magnets with other shapes could facilitate the realization of more complex circuit structures will now be investigated. Using the techniques discussed in Ch. 3.1, the necessary background is provided for the work to be discussed in Chs. 3.3, 3.4, and 3.5. Size and shape will have an effect on (a) the y-component of magnetization (which is leveraged to represent a binary 1 or 0), (b) the magnitude and direction of the field produced by a piece of magnetic material – which affects how good of a “driver” a given magnet might be, and (c) how hard or easy it is to null a magnet and/or change its polarization – which will have an effect on the overall energy required to perform computation.

First, OOMMF simulations of 4 magnets in isolation are considered. Each magnet has a different aspect ratio. For each simulation, the magnitude of $H_{clock}$ required to change
the magnetization state of each device must be determined. 60 × 90, 60 × 105, 60 × 120, and 65 × 95 nm² magnets (all 30 nm thick) were considered with the same $H_{bias}$ applied to each. Per Fig. 3.3, as a magnet’s aspect ratio increases, the magnitude of $H_{clock}$ required to switch the magnet increases as well (only two curves are included to improve graph readability). Note the “tail” in each of the curves as the magnet rotates from a ↓ orientation to an ↑ orientation. This can be explained by the fact that magnets can exhibit a slight “ringing” due to precessional switching [4] where, for example, the magnetization moves past a perfect ↑ (i.e., 90°) orientation toward the negative x-direction (e.g., -85°) before settling to a strong ↑ (90°) polarization.

This trend can be explained by the concept of the demagnetizing field. When some material is magnetized in a given direction, the field/flux lines emerge from the North pole and return at the South pole. Some of these field lines pass through the material as well. These field lines are called the demagnetizing field as they have direction opposite to that of magnetization $M$, and thus they work as a demagnetizing force. The demagnetizing field in a direction is proportional to the magnetization in that direction, $H_{demag} = N_d M$, where the demagnetizing factor, $N_d$, depends on the size and shape of the material. In general, $N_d$ is lowest along the longest dimension and highest along the shortest dimension [68]. The minimum field needed to align the magnetization of a nanomagnet along its hard axis is given by $H_{clock} = (N_x - N_y)M_s$, where $N_x$ and $N_y$ are the demagnetizing factors along the hard and easy axes of the nanomagnet respectively, and $M_s$ is the saturation magnetization. As the aspect ratio of a nanomagnet increases, the difference between $N_x$ and $N_y$ increases and thus the minimum $H_{clock}$ needed to null the magnet increases. (Though these analytical formulas assume a rectangular magnet, they would only approximate the behavior of rounded rectangle magnets, but the trends are the same.)

Simulation data is now leveraged to show how aspect ratio affects how strong of a “driver” a magnet is. To quantify this, the magnitude of the magnetic field produced by each magnet is measured by averaging the field associated with each simulation mesh
Figure 3.3. As a magnet’s aspect ratio increases, given a similar biasing field, the magnitude of $H_{\text{clock}}$ required to null/switch the magnet increases. [61].
point 4 nm directly below a magnet’s lower edge (see insets in Fig. 3.3 for an example). Of particular importance is the y-component of magnetization ($H_{y-bias}$) because (as will be seen next), it is a good measure of how strong of a “driver” a magnet is. Sample data points are presented in Table 3.1. As a magnet’s aspect ratio (and $M_y$) increases, the y-component of the generated field increases as well. While easier to null, magnets with lower aspect ratios exhibit weaker 1s and 0s and are also weaker drivers.

For magnets with a small aspect ratio, the magnetization is not totally aligned to its easy, y-axis. As a result, it is not able to produce a strong $H_{bias}$ along the y-direction, and consequently has a reduced polarization effect on its neighbor magnet. Also, small differences in the demagnetizing factor between the easy axis and the hard axis make the magnetization prone to variations with small field differences along the hard axis.

The above discussion is important as the magnitude of $H_{bias}$ will have an impact on the required magnitude of $H_{clock}$. In all of the simulation results discussed thus far, the magnets have essentially been biased by using an artificially generated external field. But as mentioned, $H_{bias}$ really comes from the magnets themselves. As $H_{bias}$ increases, the external field required to null the magnet decreases. (One can see this trend graphically by re-examining Fig. 3.1). From a circuit design perspective, this is important as it suggests

<table>
<thead>
<tr>
<th>Magnet</th>
<th>60x90</th>
<th>60x105</th>
<th>60x120</th>
</tr>
</thead>
<tbody>
<tr>
<td>$M_{y-avg}$ (A/m)</td>
<td>$-7.71 \times 10^5$</td>
<td>$-7.901 \times 10^5$</td>
<td>$-7.97 \times 10^5$</td>
</tr>
<tr>
<td>$H_{y-Bias}$ (A/m)</td>
<td>$-2.33 \times 10^5$</td>
<td>$-2.48 \times 10^5$</td>
<td>$-2.56 \times 10^5$</td>
</tr>
</tbody>
</table>

3 This can also be seen in Fig. 3.3 As the magnet’s aspect ratio increases, its y-component of magnetization increases too.
that one can reduce the overall system energy by using lower energy drive circuitry if a stronger local biasing field can be generated with magnetic material.

Obviously, a magnet’s shape influences the magnitude and direction of the biasing field that it produces. However, how a driver and a magnet being driven are oriented with respect to one another also has an effect. To quantify the magnitude of $H_{bias}$, consider the $y$-component of magnetization of the field generated by an unclocked and unbiased $60 \times 90$ nm$^2$ magnet in isolation at two different surfaces (see Fig. 3.4). The surfaces chosen are indicative of where one would want to “see” a given $H_{bias}$ – perhaps to facilitate antiferromagnetic ordering. Averaging each set of data points, one can see that at the horizontal surface, the average $y$-component of $H_{bias}$ is 51,600 A/m and at the diagonal surface the average $y$-component of $H_{bias}$ is 40,200 A/m. These results suggest that a magnet can generate a rather strong biasing field.

In reality, the biasing field generated by a magnet will be less because it is suppressed by $H_{clock}$. This effect can be quantified by again applying an external field to the $60 \times 90$ nm$^2$ magnet in isolation. Here, no biasing field was applied and average magnitude of $H_{bias}$ as a function of $H_{clock}$ was calculated for the same surfaces in Fig. 3.4. As the magnitude of $H_{clock}$ increases, the magnitude of the biasing field generated by the magnet decreases. For example, when $H_{clock} = 20,000$ A/m, the average biasing field at the horizontal surface is reduced by nearly 25% to about 40,000 A/m. The average biasing field at the diagonal surface is reduced by about 60% to 25,000 A/m. Thus, one must consider the magnitude of $H_{clock}$ when designing circuits if a specific $H_{bias}$ is required. (Another alternative is to change the geometry of the nanomagnet to help concentrate flux – increasing the magnitude of $H_{bias}$).

Thus, an ideal device will essentially be easily hard-axis biased (with a low magnitude $H_{clock}$), but can also switch state quickly to produce a strong, local $H_{bias}$ (even if $H_{clock}$ is still applied). These ideas are leveraged to accomplish circuit-level tasks such as facilitating dataflow in a line of magnets across clock wire boundaries and to help move data.
Figure 3.4. The magnitude of the y-component of the field produced by a 60 × 90 nm² magnet is measured at two different surfaces. [61].

orthogonally to the direction of $H_{\text{clock}}$.

3.3 Clock Wire Boundaries

All of the elements in an NML circuit must be controlled locally, on chip. Before looking at corner turns and crossings, first considered is the interaction between nanomagnets and the drive circuitry in more detail. Of particular interest is how information is transferred in the nanomagnets at the clock wire boundaries (see Fig. 3.5a). A line of nanomagnets is used as a case study.

The simplest way one could envision controlling a line of nanomagnets that spans multiple clock wires is as proposed in [59] and illustrated in Fig. 3.5b. Essentially, one would run current through wire (i) to generate a field that would null the first group of nanomagnets, and the magnets in this group would switch in accordance with their input. Next, one would run current through wire (ii) to generate a field that would null the second group of nanomagnets, the first magnet in group (ii) would be biased/driven by the last magnet in group (i), and the rest of the magnets in group (ii) would switch accordingly. Wires (i) and (ii) would not be active simultaneously. This clocking scheme is tested with a short, 12
Figure 3.5. (a) Different groups of magnets will be controlled by different metal wires as discussed in Ch. 2.2; (b)-(d) Wires can be active at different times to facilitate dataflow in the nanomagnets. [61].

magnet wire segment (see insets in Fig. 3.6). The line was initialized to a logically correct, antiferromagnetically coupled ground state and the polarization of the first magnet was flipped to simulate a new input. This input magnet was unclocked. A 20,000 A/m clocking field is then applied to the next 5 magnets and a 20,000 A/m field to the next 6 magnets in succession as illustrated in Fig. 3.5b. Simulation results are illustrated graphically in inset (a) in Fig. 3.6. Note that the fourth magnet in the line is frustrated and the last 8 magnets never change state (i.e., a “stuck at” fault ensues).

Why this occurs can be explained by considering the magnetization of the 5th magnet in the line (the first magnet that does not switch) as a function of the magnetization of the 6th magnet in the line (the last magnet in the wire (i) group). More specifically, the y-component of magnetization of the 5th magnet and the x-component of magnetization of the 6th magnet are considered. Examining the “No Cycle \( M_{x-6} \) vs. \( M_{y-5} \)” curve in Fig. 3.6, the 6th magnet is never completely nulled (the saturation magnetization for supermalloy is \( 8.0 \times 10^5 \) A/m [21]). The unclocked 7th magnet remains in its initial ↑ state and essentially
provides a sufficiently strong anti-bias which impedes switching of the 6th magnet. This in turn makes it more difficult to null the 5th magnet. While the 5th magnet starts to switch, it is never completely nulled and when $H_{\text{clock}}$ returns to 0, it relaxes back to its initial state.

Given this result, the clocking methodology proposed in [59] will not allow data to flow in the nanomagnets between groups (i) and (ii) if the wires are excited in the manner summarized in Fig. 3.5b. However, the simulation results presented in Fig. 3.2 have been leveraged to develop a modified clock scheme that does allow data to flow across the wire (i)/wire (ii) boundary. Looking back to Ch. 3.1, when a nanomagnet coupled to a block of magnetic material (with the magnetization of the block parallel to the direction of $H_{\text{clock}}$) is considered, the magnitude of $H_{\text{clock}}$ required for switching to occur was cut in half.

While one cannot insert blocks of magnetic material at the clock wire boundaries, one can mimic the effects of the block by exciting both wire (i) and wire (ii) simultaneously as described by Fig. 3.5c. The result of this simulation is illustrated graphically in inset (b) of Fig. 3.6. Note that the input to the wire propagates successfully between clock wire groups. Although the signal does become frustrated again in group (ii), this is because the boundary condition of wire (ii) is not controlled. If a wire group (iii) were to be placed adjacent to wire group (ii), this problem would be eliminated (this is considered further in Ch. 4.1.1). This simulation is also summarized in Fig. 3.6. Note that when the 6th magnet is nulled, $M_y$ in the 5th magnet switches from positive to negative (“Cycle $M_{x-6}$ vs. $M_{y-5}$”). By exciting both wire groups simultaneously, (a) the anti-bias produced by the first magnet in wire group (ii) is suppressed and (b) the first magnet in group (ii) contributes to the “nulling” field that the last magnet in the first line experiences.

The downside to this clocking scheme is that it could double the energy associated with the clock – which is already anticipated to be the dominant source of energy in an NML-based system. That said, one way to mitigate excess energy consumption is to consider the time evolution of this line. If one could determine when the last magnet in a wire group is sufficiently nulled, this criterion could be used to determine when to stop current flow.
Figure 3.6. $M_y$ of the 5th magnet in the line of nanomagnets shown in the insets as a function of $M_x$ of the 6th magnet in the line. Note that the magnets in wire group (i) only switch successfully if the magnetic material at the clock wire boundary is sufficiently nulled. [61].
through wire (i). This might minimize the time that both wires are active simultaneously and lead to the clocking scheme in Fig. 3.5d.

Another alternative is to leverage magnetic material at the clock wire boundaries to facilitate dataflow between clock wire groups. Conceptually, one needs a magnet that is both easy to null and an adequate driver. Looking back to Fig. 3.3, it was reported that as a magnet’s aspect ratio approached 1, it became easier to remove the y-component of magnetization and null the magnet. While this comes at the expense of a weaker biasing field (see Table 3.1), such a shape may also be more readily manipulated by the biasing fields of the surrounding magnetic material.

To facilitate dataflow between clock wire boundaries, magnets with aspect ratios that are less than but still close to 1 are leveraged. Specifically, two $50 \times 40 \text{ nm}^2$ magnets are integrated into the 12 magnet line as shown in inset (c) of Fig. 3.6. Note that the footprint of this arrangement is approximately that of a $60 \times 90 \text{ nm}^2$ magnet. However, the easy axis of each $50 \times 40 \text{ nm}^2$ magnet is parallel to the direction of $H_{\text{clock}}$. Hence these magnets should be more easily nulled than a $60 \times 90 \text{ nm}^2$ magnet where shape anisotropy makes it more difficult to magnetize along its hard axis. Next, it must be considered whether the biasing field from a $60 \times 90 \text{ nm}^2$ magnet is enough to change the magnetization of the $50 \times 40 \text{ nm}^2$ magnets in a direction against that dictated by shape anisotropy.

To determine the magnitude of the biasing field required to orient the $50 \times 40 \text{ nm}^2$ magnet up or down, one can perform simulations similar to those in Ch. 3.2. Namely, a $50 \times 40 \text{ nm}^2$ magnet in isolation is allowed to relax, and an external field is then applied along its hard axis. Per Fig. 3.7, when the magnitude of the external field is approximately 35,000 A/m, the $50 \times 40 \text{ nm}^2$ magnet has a strong y-component of magnetization ("1 Block" curve). The magnet remains in this state as long as the external field is applied. Otherwise, the magnet would relax back to the state illustrated in the inset of Fig. 3.7 when $H_{\text{clock}} = 0$.

This simulation is repeated, but with two adjacent $50 \times 40 \text{ nm}^2$ magnets. The two
Figure 3.7. $M_y$ of two different arrangements of nanomagnets as a function of $H_{\text{clock}}$ applied along the magnets’ hard axes. [61].
magnets were initialized such that \( M_s = 800,000 \text{ A/m} \) and were allowed to relax to a ground state with no applied field (see inset in Fig. 3.7 – shape anisotropy results in an x-component of magnetization while field lines emanating from each magnet lead to the coupling shown in the inset). An external field is again applied to this arrangement of nanomagnets along their hard axes. These simulation results are also illustrated in Fig. 3.7. Note that when \( H_{\text{clock}} \) is approximately 15,000 A/m, this arrangement of nanomagnets has a strong average y-component of magnetization. Looking back to Ch. 3.2, it was noted that even in the presence of an external clocking field, the biasing field produced by a \( 60 \times 90 \text{ nm}^2 \) magnet has a magnitude similar to the 15,000 A/m field reported here. Thus, this arrangement of \( 50 \times 40 \text{ nm}^2 \) magnets could be controlled by the biasing field of another magnet in the line, but also be easily nulled.

This idea is tested by repeating the 12 cell wire simulation with the \( 50 \times 40 \text{ nm}^2 \) magnet pair at the clock wire boundaries. The clocking scheme in Fig. 3.5b was used. Simulation results are illustrated by inset (c) in Fig. 3.6. Here, magnets in both wire groups change in accordance with the new input to produce a logically correct, antiferromagnetically ordered line. Moreover, only one wire group is excited at any given time which will help minimize overall system energy. The \( 50 \times 40 \text{ nm}^2 \) magnet pair remains in the state shown in inset (c) in Fig. 3.6 even when no external field is applied as the \( 60 \times 90 \text{ nm}^2 \) magnets adjacent to the pair generate the biasing field needed to preserve antiferromagnetic ordering.

3.4 Horizontal and Vertical Dataflow

To form more computationally interesting circuits, gates and wires need to be interconnected. This necessitates more complex dataflow. Considered in this section is the problem of having a horizontal line of nanomagnets drive a signal to a vertical line of nanomagnets with the output of that vertical line driving another horizontal line (i.e., “turning corners”). If it is assumed that the direction of \( H_{\text{clock}} \) is parallel to any horizontal line then dataflow will need to be facilitated in the magnets orthogonal to the direction of the applied \( H_{\text{clock}} \).
An obvious attempt to accomplish this task is to use an arrangement of nanomagnets like that shown in Fig. 3.8a. To test this structure, the magnets were initialized to a logically correct ground state and a 20,000 A/m external field was applied. The result is illustrated in Fig. 3.8b, and the y-component of magnetization of the last magnet in the vertical line appears as a function of time in Fig. 3.9 (see “No Blocks” curve). Clearly, this structure does not work as intended. For the design in Fig. 3.8b, the output magnet should be oriented ↑, but there is little if any change to the y-component of its magnetization. These results can be explained by the simulation results discussed in Fig. 3.2. The last magnet in this line behaves similarly to that described by “iso, $H_{\text{clock}}$” in Fig. 3.2. As it is not coupled to any other magnetic material, the magnitude of $H_{\text{clock}}$ needed to facilitate switching would need to be much larger than 20,000 A/m. Moreover, given the configuration of magnets in Fig. 3.8a, the second magnet in the vertical line essentially serves as an anti-bias making it even harder to switch the last magnet.

In an attempt to design a working corner turn, the arrangement of $50 \times 40 \text{ nm}^2$ magnets (discussed at the end of Ch. 3.3) is leveraged with added blocks of magnetic material oriented in the same direction as the applied $H_{\text{clock}}$ to help switch the magnet. The new design is illustrated in Fig. 3.8c. Note that the last magnet in the top horizontal line was initially oriented down, the magnets in the vertical wire were all oriented up, and the magnet adjacent to the vertical line in the bottom horizontal line was initially down. Thus, antiferromagnetic ordering was preserved in this initial state. To test this structure, the $M_y$ of the first magnet of the top horizontal line was flipped and a 20,000 A/m external field was applied to this group of magnets. The simulation result appears in Fig. 3.8d. The magnets in this structure have all switched in accordance with the new input.

While this structure appears to function as intended, ideally, the magnet being driven in the lower horizontal line will be biased such that it represents its new, logically correct state before $H_{\text{clock}} = 0$. Otherwise, it might suggest that all dataflow orthogonal to the direction of the clock would need to take place at clock wire boundaries – a significant
Figure 3.8. Vertical dataflow Design 1: before (a), after (b); Vertical dataflow Design 2: before (c), after (d). [61].
Magnetization never changes.

Figure 3.9. $M_y$ of the output of two corner turn designs vs. time. [61].
design constraint. One can determine when the last magnet in the bottom horizontal line switches by considering $M_y$ as a function of time (see “Split Magnet” in Fig. 3.9). Note that $H_{\text{clock}}$ switches from 20,000 A/m to 0 A/m at $2.61 \times 10^{-9}$ s. The last magnet in the lower wire is already biased into its logically correct state and this structure works as intended.

3.5 Wire Crossings

Here, how one can arrange a group of magnets that might function as a wire crossing is shown. The shape of a magnet is again leveraged to represent two bits of information simultaneously. A schematic of what a wire crossing might look like appears in the insets of Fig. 3.10 where the middle magnet has a component of both inputs (at left) and transfers the polarization of both inputs to both outputs (at right).

All of the above ideas have been leveraged to design the structure that appears in the Fig. 3.10 insets (with before and after snapshots). Note that here, the input magnets, output magnets, and middle magnet have all been sized differently. While ultimately this may not be required, sizing helps to ensure that the inputs will have more influence over the device that represents both states, which will in turn have more influence over the output magnets. This helps ensure the correct time evolution of this circuit and prevents “stuck at” states. This is shown quantitatively in Fig. 3.10 – where both the magnitude and the direction of the local field between the input at bottom-left (“Bottom”) and the output at top-right (“Top”) are considered. Note that the magnitude of the field between the bottom input and the middle cell is greater than the magnitude of the field between the middle cell and the top output.

3.6 Conclusions

The methods of analysis presented in Ch. 3.1 as well as the study of magnet drive strength in Ch. 3.2 are used throughout my work, and in NML research in general. The
Figure 3.10. Magnitude and direction of field between magnets of crossings vs. time. [61].
corner turn and wire crossing designs enabled by this study are both used in the benchmarking efforts presented in Ch. 5.4.1. However, the strategy for assuring proper data propagation across clock wire boundaries (Ch. 3.3) is overly primitive, and needs to be considered in more detail.
CHAPTER 4

REALISTIC CLOCK STRUCTURES AND THEIR EFFECTS ON NML CIRCUITS

Here, I present work that marks the beginning of a simulation-based effort to accurately model the effects of realistic NML clock structures on NML circuits (Ch. 4.1). This study begins by describing problems with the preliminary models from Ch. 3.3 (in Ch. 4.1.1), while improved models are presented in Ch. 4.1.2. Next, I show how these realistic models impact logical correctness and power requirements for one NML circuit of interest: an inverter chain (Ch. 4.2). The impact of clock wire design aspects and fabrication variation is also considered (Ch. 4.3). Finally, in Ch. 4.4 I consider how multi-plane clock wires (proposed in [43]) could be used to enable multiple levels of NML circuits, how field interference can exist from clock wires in different planes, and how to mitigate these effects.

4.1 A Simulation Methodology for Realistic Clock Fields

This section illustrates how simulation methodologies presented in Ch. 3.3 for small, clocked NML ensembles do not properly capture the behavior of the larger ensembles needed for application specific tasks. Problems with these preliminary models are detailed in Ch. 4.1.1, suitable improvements are detailed in Ch. 4.1.2 and simulation parameters and assumptions (e.g., material properties, etc.) are detailed in Ch. 4.1.3.

To study the behavior of clocked nanomagnets and ultimately quantify energy and delay, two simulators are leveraged. In addition to OOMMF, Ansoft’s Maxwell 2D Solver [6] is also used. Maxwell 2D is a finite element solver for electromagnetic field simulations.
4.1.1 Problems with Preliminary Models

By examining the clock wire cross section in Fig. 2.2d, it is clear that in order to accurately simulate NML ensembles patterned on top of said clock (including over the cladding itself), one needs to accurately model the field distribution and magnitude associated with a current-carrying wire. The preliminary simulations in Ch. 3.3 – and other NML published work – where parts of an NML ensemble were assumed to be controlled by separate clock wires, assumptions about field distributions, etc. have been ideal, or overly simplistic (e.g., a clock field that terminated precisely between two magnets that are spaced 12 nm apart). Here, problems with these models are discussed in more detail and how they can be improved is shown.

Moving a signal across clock wire boundaries was briefly considered in Ch. 3.3 where an AF-ordered line of twelve, $60 \times 90 \times 30$ nm$^3$ supermalloy magnets with 12 nm spacing between devices was studied. The first magnet served as an input, and it was assumed that the next five magnets would be controlled by one clock wire/field, and the last six magnets would be controlled by another clock wire/field (see Fig. 4.1a). The goal in Ch. 3.3 was to move data from group (i) to group (ii) when only one of the controlling wires was excited. These efforts assumed completely uniform fields and perfect clock wire boundaries (i.e., the magnetic field that controlled group (i) would terminate precisely between magnet N and magnet N+1 in Fig. 4.1a).

However, if a clock field was applied to just one group at a time, the new input state would not propagate through the first wire group. The first magnet in the second group effectively drove the magnets in group (i) in the opposite direction (see Fig. 4.1b) for reasons explained in Ch. 2. The net effect is a stuck-at fault defined by the state of the first magnet in group (ii)$^1$. Recall the solution was to “split” the last magnet in each group into two different pieces of magnetic material (see Fig. 3.6) for the reasons described in 

---

$^1$In the context of the discussion at the end of Ch. 2, the magnet at the end of group (i) was (a) not sufficiently hard-axis biased and/or (b) was driven backward by the first magnet in group (ii).
Figure 4.1. (a) AF-ordered line controlled by different clock wires; (b) Improper clocking can lead to mistake states; (c) Field strength 15 nm above conductors in the x- and z-directions when the first two conductors are in a steady state. [25].

Ch.3.3 While this study does consider the need to ensure that a clocked group of magnets is not driven from multiple directions (as discussed in Ch.2), it does not capture how an NML circuit would respond to the field produced by a clad, current-carrying wire.

There are two overly-simplified assumptions that make the aforementioned work incomplete. First, the magnetic field that a given clock wire generates does not in fact terminate precisely in the approximately 8-12 nm gap between magnets. Second, these previous efforts have assumed uniform clock fields. The fallacy of these assumptions is illustrated in Fig. 4.1. Fig. 4.1c shows the field distribution (obtained via Maxwell 2D simulations) assuming three parallel, clad, current-carrying wires where only the first two wires have a current applied to them. Note that field distributions are non-uniform, and minima occur at the boundaries. The cause for these distributions is well-understood. To illustrate, consider the case where current flows through both copper wires in the positive y-direction in Fig. 2.2 (assuming the coordinate system in Fig. 4.1c). As flux leaves the rightmost copper wire, it will simultaneously be returning to the leftmost wire. Most of the flux at this point is directed in the ±z-direction as it returns through the cladding to the copper
wires. The net effect is a sharp drop in magnetic field strength ($H_x$) above the yoke (i.e., the magnetic cladding around the copper wires, used interchangeably henceforth).

Another limitation in the previous studies of clocked NML lines is that a study of two clock wire groups is not sufficient to determine if a signal can reliably cross a clock wire boundary. In a line of magnets controlled by $M$ wires, (with $M > 2$) the first magnet in the line (like the input magnet in Fig. 4.1a) is a much stronger driver of the magnets in the first group than the last magnet of the first clock wire group is to the rest of the magnets in the next group. Qualitatively, when an external field is applied to the magnets in group (ii), as magnet N+1 (in Fig. 4.1a) transitions to a metastable state (i.e., as it becomes hard-axis biased), it also couples with magnet N (its driver) which couples with magnet N-1, etc. Quantitatively, in the presence of an applied clock field, the input to group (ii) produces significantly less y-directed flux (to bias its neighbor) than the original input to group (i) – as there is only air to the original input’s left and less coupling occurs (see Fig. 4.2 and discussion in Ch. 2.4.5).

Fig. 4.2 demonstrates how the initial input can be a stronger driver of its neighboring magnet than an internal magnet in an AF-ordered line. Ultimately, with a given clock field, this weaker driver can be insufficient to influence the state of each magnet on the critical path in a given wire group, and a stuck-at fault ensues. This “worst case” must be taken into account. These effects are not modeled if only considering two adjacent lines of NML devices. As such, unless noted otherwise, simulations from this point forward (a) model more than two clock wire groups, and (b) consider non-uniform field distributions, in order to better capture the actual behavior of clocked NML circuits.

4.1.2 Improved Models

Dataflow in NML ensembles needs to be considered given the aforementioned non-uniform field distributions and magnet drive strengths. To model these effects (after a few preliminary OOMMF simulations with uniform fields to gauge the field strength required
Figure 4.2. A driver magnet to an internal clock wire group may provide less y-directed flux on a neighboring device (dashed line) than the initial input to an NML ensemble (solid line). (Plotted here is the absolute value of the y-directed flux 12 nm away from a $60 \times 90 \times 30 \text{ nm}^3$ supermalloy device across a 60 nm long line.) [25].
to re-evaluate a group of $M$ nanomagnets), Maxwell 2D was used to determine what field distributions one might expect from wires clad on 3 sides with permalloy. Furthermore, three nanomagnet groups are used instead of two, and simulations (e.g., to study signal propagation) were only considered to be successful if state associated with a new input to an AF-ordered line was correctly transferred through all the magnets associated with the first two wire groups. This would suggest that the last magnet of group (i) is a sufficiently strong driver of an equally sized group.

More specifically, all of the simulations assume clocked, AF-ordered lines of 32 magnets. There are ten magnets per clock wire group and one input magnet. One magnet was also placed after the last clock wire group to mimic a fourth clock wire group. The number of devices was fixed in this study (a) to allow fine-grained simulations to finish in a reasonable time frame (approximately 1.5 days per simulation on a high-performance computing cluster, e.g., a 2.6 GHz AMD Opteron with 4 cores), and (b) in order to consider how the power of a well-behaved circuit is affected by magnet and clock wire scaling. For this study, only AF-ordered lines were considered. Regarding the issue of boundary conditions, if necessary, one could always cross boundaries with lines, and control other logic elements by the more uniform fields in the middle of a clock wire. Regarding the issue of performance projections (to be discussed in Ch. 5.4), line simulations are used to estimate critical path delays in larger ensembles.

4.1.3 Simulation Parameters and Assumptions

The Maxwell 2D finite element solver is used to generate field distributions for the clock wires; in doing so, multiple simulations are run with different current magnitudes to determine what current is needed to generate a given clock field. Next, the field data is converted from Maxwell 2D for use in OOMMF simulations to determine how magnet ensembles respond given the more realistic field distributions one should expect. Unless otherwise noted, all of the simulations assume magnets made from supermalloy (see
Ch. 3.1.2 for parameters), but with a damping coefficient of 0.1 instead of 0.5 which is more in-line with experimental results [27] (0.1 is thus used throughout this work unless otherwise noted). Supermalloy is also used in all ongoing experimental efforts.

A copper wire is assumed, and it is always surrounded by 10 nm of oxide for purposes of electrical insulation (see Fig. 2.2d). This is a reflection of early fabrication efforts. Recently, this oxide has been done away with as it was found to be unnecessary. However, all clock wires in this work consider wires with this oxide (note there is no functional difference). The width of the copper wire is selected to accommodate 10 magnets (after accounting for the cladding width). Thus, as magnet dimensions scale down, wire width will decrease. For all simulations, clock wire thickness was fixed at 600 nm unless otherwise noted. Regarding wire length, for power estimates, it will always be assumed that a wire must accommodate 25 parallel lines of magnets. Thus, any drive circuitry will have to support the same NML “function”, so clock wire length will vary in power estimates. Finally, note that a small, out-of-plane component of the field ($H_z$) is also modeled (see Fig. 4.1c). However, it has a minimal effect on the outcome of simulations (as the z-axis is a “super hard” axis).

4.2 Logically Correct Behavior with Realistic Boundary Conditions

While the 2-phase clock model in Ch. 3.3 is desirable from the standpoint of lower power, the magnets controlled by such a clock do not consistently settle into a logically correct state. As such, the 4-phase alternative also proposed in Ch. 3.3 is revisited. Namely, wires that control magnets in groups (i) and (ii) would be turned on at the same time (see inset in Fig. 4.1c). Here, in order to change the state of the magnets in group (i), they should only see one strong driver (in this case, the input magnet). After all magnets in group (i)

---

2Because a fixed clock wire thickness could lead to disproportionately scaled wire resistances, power estimates were also calculated for square aspect ratio wires. However, given the geometries considered, the impact on total power was minor and this data is not reported.
are biased in a logically correct state, one could remove the clock field from group (i) and begin to pass current through the clock wire that controls group (iii). The last magnet in group (i) would then drive group (ii). The disadvantage to this scheme is that the power associated with the clock wires will be increased and throughput will be decreased. While this is undesirable, lower power operation is of no use if a circuit’s output is not logically correct.

4.2.1 Perfect vs. Realistic Boundaries

Before applying the realistic clock fields, lines of magnets controlled with a 4-phase clock and perfect boundary conditions were first simulated. (Perfect boundary conditions were considered first as it allows quantification of whether ensuring logically correct behavior will result in increased system-level power.) With this method, for all magnet sizes, a new input state was able to propagate through the magnets in the second clock wire group. Table 4.1 lists fields required to facilitate switching in each line, the current needed to generate said fields for the reported clock wire geometry, and the power associated with a wire switching event. (Note that for the $60 \times 90 \times 30 \text{ nm}^3$, realistic case, only groups of 8 magnets could be well controlled.)

When the simulations discussed above were repeated with nonuniform and discontinuous fields (magnets were centered directly over the clock wire boundary) the 4-phase clock again facilitated logically correct switching past the second wire group. However, logical correctness comes at the expense of higher power as a larger field is needed for logically correct re-evaluation (see Table 4.1). If a field with the same strength as the perfect clock case is used, stuck-at faults occur (see the lower curve in Fig. 4.3a). The stronger field is necessary as the last magnet in each clock group is not as well controlled due to the drop in field strength over the yoke. Per Fig. 4.3a, a stronger external field is able to compensate for the lower field at the boundary and to remove the previous $M_y$. 

48
TABLE 4.1

POWER PER WIRE ASSUMING 25 PARALLEL LINES ATOP EACH CLOCK WIRE.

<table>
<thead>
<tr>
<th>Magnet size (nm^3)</th>
<th>Clock Scheme</th>
<th>Field to Switch (A/m)</th>
<th>Current to Produce Field (mA)</th>
<th>Length (nm)</th>
<th>Power (µW)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 × 60 × 20</td>
<td>Perfect</td>
<td>1989</td>
<td>1.26</td>
<td>3060</td>
<td>0.177</td>
</tr>
<tr>
<td>40 × 60 × 20</td>
<td>Realistic</td>
<td>3979</td>
<td>2.29</td>
<td>3060</td>
<td>0.586</td>
</tr>
<tr>
<td>50 × 75 × 25</td>
<td>Perfect</td>
<td>9947</td>
<td>6.80</td>
<td>3825</td>
<td>5.00</td>
</tr>
<tr>
<td>50 × 75 × 25</td>
<td>Realistic</td>
<td>15915</td>
<td>10.87</td>
<td>3825</td>
<td>12.80</td>
</tr>
<tr>
<td>60 × 90 × 30</td>
<td>Perfect</td>
<td>11937</td>
<td>9.34</td>
<td>4590</td>
<td>9.23</td>
</tr>
<tr>
<td>60 × 90 × 30</td>
<td>Realistic</td>
<td>15915</td>
<td>10.87</td>
<td>3825</td>
<td>12.80</td>
</tr>
</tbody>
</table>

Figure 4.3. (a) For realistic clock distributions, higher external fields are required to facilitate a logically correct state transitions at the second-third clock wire boundary; (b) In an annealed yoke, field lines should not upset magnetic device state. [24].
4.2.2 Effects of the Clock Wire Cladding

Whether or not the clock wire yoke itself could interfere with signal state is now considered further. The clock lines are similar to MRAM word and bit lines. In commercial MRAM chips, the sidewalls are annealed such that the long axis is a true easy axis when no current flows through the copper [69]. Annealing helps to ensure that no domains form, which could affect the field produced while current is flowing through a wire and also lead to fields at the yoke’s surface that could alter the logical state of a magnet when wires are not excited.

Of most concern is whether strong magnetic fields from the yoke material could work to alter the state of a driver nanomagnet. To study this problem, the yoke sidewall was modeled in OOMMF as a $25 \times 500 \times 4000$ nm$^3$ piece of permalloy magnetized along its hard axis (see inset in Fig. 4.3b). To ensure that no domains form in the material in these simulations, a 120,000 A/m field was applied directly to the magnetic material (but not to the space around it). This mimics the effects of annealing. The x-, y-, and z-components of any magnetic field were then measured 30 nm above the length of the sidewall; all are illustrated in Fig. 4.3b. The only stray fields present that might unduly alter a magnet’s state are at the very far ends of the wire. Also, hysteresis in the yoke should not affect a device’s state.

4.3 Clock Wire Design and Fabrication Variations

Clock wire cladding is indispensable for generating strong clocking fields with reasonable currents. Cladding thickness, depth, and relative positions to magnets may impact the correct functionality of a circuit element. These design aspects are investigated in this section.
4.3.1 Yoke Variation

As evidenced by the Maxwell 2D simulation results captured in Fig. 4.3, as yoke thickness increases, field discontinuities increase as well – which could be a source of faulty behavior if suitably high external/clock fields are not employed. While this seemingly suggests that yoke sidewalls should be as narrow as possible, high-aspect ratio sidewalls will be more difficult to fabricate and a minimum cladding thickness is needed [69].

The simulation results in Table 4.1 assumed 50 nm of cladding between copper wires. Further simulations to determine if data would propagate successfully through lines of $40 \times 60 \times 20$ nm$^3$ magnets with $30$ nm of cladding show lines were well controlled with similar clocking fields (e.g. 3,183 A/m) as discontinuities decrease (see Fig. 4.1c). However, if a minimum yoke thickness is required (for practical fabrication or flux concentration), and magnet sizes are further scaled, multiple magnets may reside over a boundary. To study the impact on logical correctness, the same lines of $40 \times 60 \times 20$ nm$^3$ magnets were considered, and $100$ nm of cladding between copper wires was assumed. Two magnets now sit over a clock wire boundary. Larger external fields (7,957 A/m) were needed to ensure both magnets were placed in a sufficient metastable state to avoid stuck-at faults.

The results in Table 4.1 suggest that as magnet size decreases, the field magnitude required to facilitate switching decreases as well. This is actually not the case, and this trend can be attributed to scaling the spacing between the magnets, which increases device coupling. This trend will be studied in detail in Ch. 5.3.2 while Ch. 5.3.4 presents trends associated with scaling magnet sizes. The practical impact of scaling magnet dimensions is that with sufficiently small magnet sizes, yoke thickness may impede these trends as more than one magnet will be less well-controlled.

4.3.2 Wire Thickness

The fabrication process discussed in Ch. 2 should lend itself to thick wires with high aspect ratio sidewalls. Clock wire thickness is important as thicker wires have lower re-
Figure 4.4. Magnetic field orthogonal to the direction of current flow along the surface of a 2 micron wide wire (assuming 4 different wire thicknesses). [24].
sistance, decreasing the overall energy associated with clocking. A previous study \[59\] reported that power consumption due to square waves and sine waves are both upper bounded by \(I^2R/2\). Thus, reducing \(R\) reduces \(P\). As the permeability of the yoke is so much larger than that of air, most of the reluctance (the magnetic equivalent of resistance) of the air-above-wire and yoke-around-wire magnetic field-line circuit is contained in the air portion above the wire. Increasing the thickness of the wire (and yoke) adds little to total reluctance.

To study further, a 2000 nm wide copper clock wire was considered with Ansoft's Maxwell 2D simulator. The wire was clad on 3 sides with 50 nm of permalloy. The expected field along the unclad surface of the wire – assuming four different thicknesses and a 16 mA current – is illustrated in Fig. 4.4. Per the inset, changing wire thickness by 400 nm only changes the field magnitude at the surface by approximately 40 A/m – or about 0.6%. Thus, if fabricatable, the thickness of a clad wire should help to reduce power and should not affect logical correctness – as the same magnitude clocking field will be produced.

4.3.3 Magnet Misalignment

This section is concluded by considering whether magnet-clock wire boundary misalignment could induce faulty behavior. For the results discussed in Ch. 4.2.1 a magnet is assumed to be centered directly over a clock wire boundary. Fabrication tolerances will not be so precise. That said, simulations were performed where the lines of magnets are shifted to the left or to the right by 24 nm (half the width of a magnet plus spacing between adjacent magnets) to mimic worst case fabrication misalignment. In each instance, the magnet at the boundary was placed in a metastable state such that it could be biased into a new, logically correct state by its neighbor. No higher fields were required to avoid faulty behavior.
4.4 Computing with Multiple Planes of NML via Multi-plane Clocking

[43] proposes a new approach that uses multiple planes of clock wires to implement the line clock in order to lower energy consumption and simplify fabrication. [65] briefly describes how multiple planes of NML could be controlled via multiple planes of clock wires. Here, I describe how multi-plane clocking could be leveraged to enable computing with multiple planes of NML. One key question needs to be addressed: can the multi-plane clocks generate field distributions capable of clocking NML circuits in multiple planes? There are two major aspects that must be considered in order to answer this question. First, to enable multiple planes of NML, one must leverage a multi-plane clock where one plane of clock wires will have cladding on only two sides (versus three sides in the other planes, as in Ch. 2.2, etc.). Thus, I quantify the effects on field distribution resulting from the removal of this cladding. Second, in Ch. [4.2.2] I demonstrated how field interference in a single plane of line clocks (as in Fig. 2.2d, Fig. 4.1c) should not be strong enough to affect device state. The multi-plane clocking approach raises the question of whether field interference could exist between the multiple planes of clock wires and potentially upset NML state in other levels. Thus, I also describe my efforts to study the effects of field interference in multi-plane clocks.

In Ch. [4.4.1] I discuss background on multi-plane clocking as proposed in [43]. This includes my initial simulation efforts to characterize fields produced by the clock wires. Next, in Ch. [4.4.2] I present results showing how clock wires with cladding on two sides result in lower field distributions, how field interference can be a problem when considering multiple planes of NML circuits, and how to account for these factors to generate a nearly uniform field distribution over the planes of NML circuits. I conclude in Ch. [4.4.2.5] with

---

[^3]: The multi-plane clocking approach was proposed by Steven Kurtz [43]. I carried out initial simulation work to characterize field distributions and describe these efforts in this work. In [43], he considers two planes of clock wires that control a single plane of NML circuits. He considers reduced energy, simplified fabrication, etc. in this context. In this work, I focus on leveraging greater than two planes of wires to control multiple planes of NML circuits, and consider the effects of clock wires with cladding on two sides, and the effects of field interference between these multiple planes of NML devices.
a discussion of the impacts on current/power requirements, and tradeoffs between wires with different aspect ratios.

4.4.1 Multi-plane Clocking

[43] describes a new method to implement the current-carrying line clocking approach. In this method, multiple planes of clock wires are fabricated above and below a single plane of NML circuits. As an example, Fig. 4.5 shows two clock wires controlling a single plane of NML devices. The multi-plane approach offers a number of potential benefits as compared to the single-plane approach. Per [43], it should reduce fabrication complexity, routing overhead, and inter-wire capacitance. Also, as will be seen in Ch. 4.4.1.1, it could mitigate the drop in field strength on magnets patterned above the magnetic cladding around each wire (which resulted in increased current and energy as seen in Ch. 4.2.1).

[43] proposes controlling a single plane of nanomagnets with the multi-plane clock structures, but with more planes of clock wires one could control multiple planes of NML circuits (described briefly in [65]). For example, consider three levels of NML clock wires as in Fig. 4.6 where the middlemost clock wire has cladding only on two sides. One potential benefit of this method is that one can achieve twice the logic density for the same chip area. Theoretically, one could increase the number of clock wire levels further and realize more than two planes of NML circuits. However, each further level increases the routing complexity – this will need to be studied closely. Furthermore, it is not currently feasible to build these structures experimentally, as the typical approach to read circuit state via MFM will no longer be possible (as the higher clock levels obstruct the MFM tip). In order to study this structure via experiment, NML magnetic-to-electrical interfaces are needed, which are still a subject of ongoing study [53].

Key questions in this study are whether the field from the center wire will be strong enough to control NML ensembles (if not, increased current may be required) and whether stray fields from clock wires controlling one plane of NML devices can adversely affect
Figure 4.5. Multi-plane clocking with a single plane of nanomagnets. Copper is gray, magnetic cladding is black. Dashed vertical line indicates measurement point for overlap of two wires.
the state of magnets in other planes. Before considering these issues, simulation work is discussed that characterizes fields generated by two planes of clock wires that laid the foundation for the study of more planes.

4.4.1.1 Field Distributions

In this section, initial simulation efforts to determine the field distributions created by two levels of multi-plane wires are described.

Similar to the work in Ch. 4.1.2 whether clock wires in multiple planes can generate the fields required to clock NML ensembles needs to be determined. There are two new critical design parameters when considering clock wires in multiple planes: the spacing between vertically-adjacent wires, and the spacing between horizontally-adjacent wires. Fig. 4.5 depicts an initial design to demonstrate the field distributions produced by these clock wires. The dashed vertical line indicates the points on the two wires at which horizontal spacing between adjacent clock wires is measured. As an initial example, consider a Maxwell simulation of two 500 $\times$ 500 nm$^2$ copper wires clad with 50 nm of permalloy on three sides. As in Ch. 4.1.3 a 2-D simulation is used, i.e., the length of the wires is assumed to be infinite. The vertical spacing between the two wires is 80 nm, and the horizontal spacing is 0 nm (i.e., the right-side cladding of the lower wire is directly in line with the left-side cladding of the top wire). A 5 mA current is placed on the lower wire in the negative x-direction (into the page), and on the upper wire in the positive x-direction (out of the page). Thus, the two wires produce fields from left-to-right over the magnets patterned between them. Design 1 (red curve) in Fig. 4.7 shows the field distribution produced by this setup. Note that there is a drop in magnetic field strength in the region where the cladding overlaps similar to that in Fig. 4.1c for single-plane wires. As in the case of the single plane wires, the fields above the cladding are concentrated strongly in the z-direction as they enter/exit the cladding, reducing the x-directed field.

Now consider what happens if the overlap and/or vertical spacing between the two
Figure 4.6. Multi-plane clocking with two planes of nanomagnets. Red Xs, circles denote direction in which current is applied but not any specific excitation pattern.
clock wires are changed. Initially considered were the effects of changing horizontal overlap, followed by the effects of changing vertical overlap. Table 4.2 lists each of the designs considered, and field distributions for each are presented in the same plot in Fig. 4.7.

Due to the drop in field strength in Design 1, considered first are the effects of increasing the horizontal spacing. Designs 2-4 represent three different horizontal spacings (40, 100, 180 nm) with vertical spacing fixed at 80 nm. As horizontal spacing increases, there is more overlap between the two clock wires and the fields generated by the two wires add together (see Fig. 4.5). The spike in field strength could, for example, be leveraged to clock the high aspect ratio latch structures that will be considered in Ch. 6.

Next, consider what happens as the vertical spacing between the two clock wires is increased. Designs 5-7 represent three different vertical spacings (120, 160, 200 nm) and a fixed horizontal spacing of 40 nm. As shown in Fig. 4.5 as the wires are moved farther apart, the spike due to clock wire overlap begins to smooth out. Thus, Design 7 produces a clock field over the magnets that approaches the uniform fields discussed in Ch. 4.1.1. This should result in decreased field, and thus current requirements, as increased current/field to mitigate the sharp field strength drop due to cladding are no longer required. Therefore, in the forthcoming discussion on field interference, Design 7 will be considered (i.e., 40 nm horizontal spacing, 200 nm vertical spacing between clock wires).

4.4.2 Field Distributions in Multiple Planes of NML

Now discussed are challenges arising from a design that leverages more than two planes of clock wires to enable multiple planes of NML circuits. Namely, the effects of removing the cladding on the top of the magnets in Clock Wire Level 2 (see Fig. 4.6) are quantified, as well as the effects of potential field interference between levels.
### TABLE 4.2

SUMMARY OF DESIGNS USED TO STUDY THE EFFECTS OF CHANGING HORIZONTAL AND VERTICAL SPACING BETWEEN TWO MULTI-PLANE CLOCK WIRES AS IN FIG. 4.5.

<table>
<thead>
<tr>
<th>Design</th>
<th>Vertical Spacing</th>
<th>Horizontal Spacing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>80 nm</td>
<td>0 nm</td>
</tr>
<tr>
<td>2</td>
<td>80 nm</td>
<td>40 nm</td>
</tr>
<tr>
<td>3</td>
<td>80 nm</td>
<td>100 nm</td>
</tr>
<tr>
<td>4</td>
<td>80 nm</td>
<td>180 nm</td>
</tr>
<tr>
<td>5</td>
<td>120 nm</td>
<td>40 nm</td>
</tr>
<tr>
<td>6</td>
<td>160 nm</td>
<td>40 nm</td>
</tr>
<tr>
<td>7</td>
<td>200 nm</td>
<td>40 nm</td>
</tr>
</tbody>
</table>
Figure 4.7. Field strength as a function of position for the seven initial multi-plane clocking designs. Note that line markers do not denote data points (data was gathered every 1 nm in the x-direction) but rather distinguish different lines.

4.4.2.1 Origins of Potential Field Interference

Ch. 4.2.2 shows how out-of-plane magnetic fields from the cladding surrounding current-carrying clock wires won’t adversely affect the state of NML devices. Micromagnetic simulation of the cladding sidewall showed that the only stray fields that could possibly upset device state are at the ends of the long wire (easily avoided by simply not placing NML devices near these edges of the wires). Furthermore, simulation results for adjacent clock wires in a single plane show that fields are confined in the x-direction. This is illustrated in Fig. 4.1, which depicts the field over three adjacent clock wires when the first two wires are excited. Note that the field above the third clock wire (i.e., the wire with no current applied, marked “Off” in the plot) is insignificant (between -2,000 and 1,000 A/m) compared to the minimum field required to clock the $40 \times 60 \times 20 \, \text{nm}^3$ magnets (i.e., 3,979 A/m). Thus, in the single-plane clocking approach, there should be no danger of a clock wire adversely affecting the state of NML devices on other wires that are in the “driving”
state (i.e., with no clock field applied).

In a multi-plane clock [43, 65] as in Fig. 4.5 or Fig. 4.6, the field from the cladding will exhibit the same behavior as in Ch. 4.2.2. Horizontally-adjacent clock wires should also behave in the same way as single-plane clock wires and have no effect on devices in other clock groups. Finally, as in Ch. 4.1.3, the field in the z-dimension still will not be strong enough to affect device state (recall that the z-dimension is a “super hard” axis).

However, in the case where multiple planes of NML devices are controlled by more than two levels of clock wires (e.g., as in Fig. 4.6), the x-directed field produced by wires in one level could impact devices in other vertically-adjacent levels. The potential for field interference will likely depend on the distance between the planes of NML circuits, as field strength diminishes with distance from the clock wire (as seen in previous work, including [2]).

For example, consider the direction in which fields are applied in the structure depicted in Fig. 4.6 (red Xs, circles denote direction current is applied, but not necessarily any specific excitation pattern). The clock wires in Clock Level 1 will generate fields in the positive x-direction, while clock wires in Clock Level 3 will generate fields in the negative x-direction. Finally, as current is applied in the opposite direction in Clock Level 2, that clock wire will generate fields on NML Level 1 in the positive x-direction, and on NML Level 2 in the negative x-direction. (Recall that signal propagation in the NML circuits will be in the same direction, i.e., left-to-right.) In this configuration, the fields generated by the clock wires in Clock Level 1 will, for example, extend to NML Level 2, and the strength of this field will depend on distance from the wire and current applied to the wire. Thus, the fields in the negative x-direction that are applied to clock the magnets in NML Level 2 would be reduced due to the effect of field interference from Clock Level 1. Similarly, the fields applied to NML Level 1 could be reduced in the same way due to interference from Clock Level 3. To what extent the field interference will affect field strength, and thus circuit operation, must now be determined. Any necessary methods to mitigate this
4.4.2.2 Simulation Methodology

In this study, three levels of clock wires that could control two planes of NML circuits are considered (as depicted in Fig. 4.6). Future work could consider simulations of greater than three levels (e.g., four were discussed in [65] but not simulated), but the observations and results from this study should apply in those cases.

As discussed in Ch. 4.3.2 and [3], a high aspect ratio (thickness/width of the copper) is desired for the wire geometries. Increasing the aspect ratio of the wires should lead to lower resistance, and per Ch. 4.3.2 wires with different aspect ratios result in a very small change in the clock field produced (e.g., with a 400 nm change in thickness, only a 0.6% change in field strength). For this reason, [3] considers aspect ratios between approximately 1 to 2, and in Ch. 5.4.1 an aspect ratio of 3 is assumed. Thus, when considering the multi-plane clocking approach, an aspect ratio of 1 and an aspect ratio of 3 will be considered (in this work, it is assumed that the aspect ratio is measured for the copper wire only).

The clock wires are composed of copper (500 nm × 500 nm, and 500 × 1500 nm) and the thickness of the surrounding permalloy cladding is fixed at 50 nm for this study. Vertical spacing between clock wires is 200 nm, and a horizontal overlap between adjacent wires of 40 nm is considered. This corresponds with Design 7 in Table 4.2 and Fig. 4.7 which, as discussed in Ch. 4.4.1.1 results in a smooth field distribution over the NML devices. Simulations are again carried out using Ansoft’s Maxwell 2D solver [6]. A current of 5 mA is placed on all clock wires. The current is oriented in the positive direction in the center clock wire, and in the negative direction in the other four wires. Thus, the clock field in the lower plane of magnets is applied in the positive x-direction, and the clock field in the upper plane of magnets is applied in the negative x-direction. Note that data propagates in the positive x-direction in both planes of magnets (which are not included in the Maxwell simulations). Per Ch. 3.1.2 field application in either direction should
facilitate propagation, and the direction of propagation is controlled by the 3-phase clock described in Ch. 5.1.1.

4.4.2.3 Results

This section describes results of simulations to understand the effects of field interference. Also discussed are the effects on field distribution from removing the cladding from the center wires (e.g., the wire in Clock Level 2).

First considered are clock wires with an aspect ratio of 1.0. As expected, the clock wires produce fields similar in magnitude and shape to those in Fig. 4.7. For example, consider Fig. 4.8 which depicts the field in NML Level 1 when only the bottom left and center clock wires are excited. The field is measured 100 nm above the wire where NML devices would be placed (the measurement is taken over the width of the entire structure – see inset, red line in Fig. 4.8). The average field over the first clock group in NML Level 1 is 6.0 mT. However, the average field over the second clock group is only 3.7 mT. The majority of the field in that region is due to the contribution from the clock wire in Clock Level 2 (i.e., the wire with cladding only on two sides). Note that the other wires in the multi-level arrangement produce fields of similar magnitude and shape as the design is symmetric. The reduction in the second group is due to the removal of the cladding on the top of the wire in Clock Level 2. Therefore, increased current would be required in these wires in order to enable signal propagation in the nanomagnets. This approach will be discussed and the required increased current quantified in Ch. 4.4.2.5.

Next, consider the effects of field interference from clock wires in Clock Level 1 on the devices in NML Level 2 for wires with an aspect ratio of 1.0. Note that since the structures are symmetric, the result should be the same as if the effect of Clock Wire Level 3 on NML Level 1 were considered. Fig. 4.9 depicts the $H_x$ field distributions for one of the possible excitation patterns of the five clock wires per the 3-phase clock (to be discussed in Ch. 5.1.1). The first group is switching, the second group is in the buffer state (i.e., field
Figure 4.8. Fields generated when the lower left and center wires are excited (see inset, red X). Aspect ratio of all wires is 1.0. Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
is applied as with the 4-phase clock to stop undesirable back propagation), and the last is 
*driving*, i.e., the clock is removed. The expected average field seen in NML Level 2 is
-6.0 mT in the first (switching) group, and -3.7 mT in the second (buffer) group (recall the
center wire produces lower field for the same current). However, what is seen in this
case is a field of -5.6 mT in the first group, and -2.9 mT in the driving group – the field
magnitude is thus effectively reduced by 0.4 mT and 0.8 mT in the first and second groups,
respectively. (For ease of comparison, Fig. 4.10 plots the absolute value of the curves in
Fig. 4.8 and Fig. 4.9 together.)

The reduction in field strength in NML Level 2 can be explained by considering the
fields present in NML Level 2 when only the leftmost wire in NML Level 1 is excited.
This field distribution is depicted in Fig. 4.11. This wire generated an average field of 0.65
mT over the first group in NML Level 2, and 0.92 mT over the second group in the same
level. Since these fields are in the positive x-direction, they will cancel with the fields from
the clock wires in Clock Wire Level 3 and reduce the field strength seen by the magnets in
the first two groups. The effect is similar when considering the effect of Clock Level 3 on
NML Level 1, and for other excitation patterns. This reduced field strength could lead to
logically incorrect operation in the clock wires. (Ch. 4.4.2.5 discusses methods to mitigate
these undesirable results.)

Consider next the case where the aspect ratio of the copper wire is raised to 3.0 (again,
desired per [3] in terms of lowered resistance). First, it is determined whether these wires
produce similar field distributions as in the single plane case. Fig. 4.12 depicts the field
generated by the leftmost clock wire in Clock Level 1 and the center clock wire in Clock
Level 2 on NML Level 1. Similar to the wires with AR 1.0, the field generated is on average
5.6 mT in the first group of NML Level 1, and 3.5 mT in the second group. Also, note that
again the field drops to between 0 and -0.35 mT over the neighboring clock wires. The
other clock wires generate fields of similar shape and magnitude. As expected, the average
field strength generated by the center clock wire is again less than that produced by the
Figure 4.9. Fields generated in NML Level 2. Aspect ratio of all wires is 1.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
other clock wires, as the cladding is removed from the top of the clock wire. However, the effect is more pronounced than in the case of the wires with AR of 1.0, as even less flux is directed in the x-direction. Therefore, more energy will likely have to be expended in order to enable signal propagation by these clock wires. The impacts of this drop in field strength will be discussed in Ch. 4.4.2.5.

Next, the effects of field interference for clock wires with an aspect ratio of 3.0 are considered. Fig. 4.13 depicts the field distribution in NML Level 2 for the excitation pattern described above (first group switching, second buffer, third driving). As with the clock wires with aspect ratio 1.0, the fields produced in the first (switching) and second (buffer) groups are reduced. In this case, the average field in the first group is -5.1 mT and -2.8 mT in the second group. The field is thus 0.5 mT lower than expected in the first group, and 0.7 mT lower than expected in the second group. (For ease of comparison, Fig. 4.14 plots the absolute value of the curves in Fig. 4.12 and Fig. 4.13 together.) Once again, this can
Figure 4.11. Fields generated when the lower left wire is excited (see inset, red X). Aspect ratio of all wires is 1.0. Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
Figure 4.12. Fields generated in NML Level 1 when the lower left and center clock wires (aspect ratio of all wires is 3.0) are excited (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
be attributed to the effects from the leftmost clock wire in Clock Wire Level 1. Fig. 4.15 shows the impact of this clock wire on NML Level 2. An average of 0.11 mT is seen in the first group in NML Level 2 when only the leftmost wire in Clock Wire Level 1 is excited. A much stronger average field of 0.50 mT is generated over the second clock group in NML Level 2. The field interference in this case is lower than for the wires with aspect ratio 1.0. Table 4.3 summarizes the results gathered for both aspect ratios. Note that data in Table 4.3 is average of fields in first and second clock groups (G1, G2 respectively) in NML Level 2.

4.4.2.4 Solution

The reduction in field strength due to the interference between clock wire levels could cause signal propagation to fail. In order to mitigate this issue – as well as the reduced field produced by the center clock wires (those with cladding on only two sides) – increased current will be required in all clock wires. As an example, consider Fig. 4.16 which considers the wires with aspect ratio of 1.0 and an increased current: from 5 mA to 6 mA in Clock Wire Levels 1/3, and from 5 mA to 11 mA in Clock Wire Level 2. With these increased currents, one can achieve a nearly smooth field distribution at an average of -6.9 mT in the

TABLE 4.3

SUMMARY OF FIELD DISTRIBUTIONS FOR CLOCK WIRES WITH ARS OF 1.0 AND 3.0.

<table>
<thead>
<tr>
<th>Aspect Ratio</th>
<th>Expected G1</th>
<th>Expected G2</th>
<th>Actual G1</th>
<th>Actual G2</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>6.0 mT</td>
<td>3.7 mT</td>
<td>5.6 mT</td>
<td>2.9 mT</td>
<td>0.4 mT, 0.8 mT</td>
</tr>
<tr>
<td>3.0</td>
<td>5.6 mT</td>
<td>3.5 mT</td>
<td>5.1 mT</td>
<td>2.8 mT</td>
<td>0.5 to 0.7 mT</td>
</tr>
</tbody>
</table>
Figure 4.13. Fields generated in NML Level 2. Aspect ratio of all wires is 3.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
first clock group, and -6.7 mT in the second clock group. Note the field reaches -9.5 mT in one region, which will not be detrimental (but if desired, could possibly be smoothed out by further adjusting the horizontal overlap between clock wires). Thus, the field interference as well as the weakened fields produced by the clock wires in Clock Wire Level 2 have been taken into account. To compare to the case where clock wires have an aspect ratio of 3.0, results are presented using the same currents as above. Thus, rather than tuning current to achieve the same field strength, identical currents are used and the field distributions are compared. Results are presented in Fig. 4.17 and show that the field distribution is again nearly smooth; the average field is -6.8 mT in the first clock group, and -6.5 mT in the second clock group.
Figure 4.15. Fields generated in NML Level 2 when the lower left clock wire (aspect ratio of all wires is 3.0) is excited (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
Figure 4.16. Fields generated in NML Level 2 for currents of 6 mA in all but the center wire where current is 11 mA. Aspect ratio of all wires is 1.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
Figure 4.17. Fields generated in NML Level 2 for currents of 6 mA in all but the center wire where current is 11 mA. Aspect ratio of all wires is 3.0. Excitation pattern per 3-phase clock when first group is driving (see inset, red X). Field is measured 100 nm above the wire along the width of the multi-level structure (see inset, red line). Note that there are no NML devices in this simulation, the devices pictured here are simply to show where they would be placed.
4.4.2.5 Summary

How to generate uniform field distributions in multiple planes of NML circuits using multi-plane clocking was described. Increased current is required to achieve the uniform distribution mainly due to reduced field strengths in Clock Wire Level 2, but also in part from the effects of field interference between Clock Wire Levels 1 and 3. The current required in Clock Wire Level 2 is almost twice that required in Clock Wire Level 1 to achieve the same, uniform field. Future work should consider different wire geometries and vertical/horizontal spacings between wire levels in order to reduce this requirement (if possible).

The effects of increasing the aspect ratio of the wires from 1.0 to 3.0 was also considered. Notably, though the field interference between levels is reduced when aspect ratio is increased, the fields generated by these wires are slightly weaker, so the same current is required to achieve comparable uniform field distributions to the low aspect ratio case. However, as described in [3], high aspect ratio wires would lead to decreased resistance and thus could require less power to generate these fields. Future work should consider a large design space sweep that includes multiple aspect ratios in this context, as well as in the context of the work in [43] which suggests that power savings from high aspect ratio wires might be reduced due to increased inter-wire capacitance. Finally, it will be unfeasible to study multiple planes of NML in experiment until NML circuits and clock wires can be integrated with I/O, as the state of the magnets will no longer be able to be probed with MFM.

4.5 Conclusions

The simulation methodology presented in Ch. 4.1 will be used in all future work that considers realistic clock structures. The study of logical correctness and power gives guidelines on power requirements for circuit design with NML clock lines, and forms the basis
for an extended study in Ch. 5. The study in Ch. 4.3 demonstrates how NML circuits are robust in the face of cladding variation, clock wire thickness, and magnet misalignment over the cladding. Finally, the study in Ch. 4.4 shows how multiple planes of clock wires could enable multiple planes of NML circuits. There will be a tradeoff between the increase in energy consumption (due to field interference and decreased field strength from clock wires with cladding on only two sides), and the increased logic density achieved for the same area. The tradeoff will need to be taken into account in the context of any application of interest.

Going forward, NML study can focus on circuit and architecture design with a clear understanding of increased field requirements that will ensure logical correctness. Further study beyond power and correctness is desired to fully characterize NML performance and to prove how NML can satisfy the digital logic tenet for uni-directional dataflow.
CHAPTER 5

CLOCKING THEORY TO QUANTIFY ENERGY AND DELAY AND TO ENSURE DATAFLOW UNIDIRECTIONALITY

As an extension of the study of realistic clock wires, I describe new clocking theory to quantify energy and delay for NML circuits in general (Ch. 5.1) as well as a study of three clocking methods (Ch. 5.2). This study demonstrates how the clock can impose strict dataflow directionality and satisfy the relevant tenet for digital logic devices discussed in Ch. 2.4.5. The studies in previous work considered only logical correctness and the power consumed in the clock wires. This new clocking theory allows for a more detailed study of NML inverter chains (Ch. 5.3) that also considers energy and delay. Furthermore, it enables energy and delay benchmarking and a comparison to a CMOS ripple carry adder (Ch. 5.4), which is used to demonstrate how NML can outperform low-power CMOS in terms of the energy delay product benchmark. Alternative paths to lower power are also considered in Ch. 5.5 and Ch. 5.6.

5.1 Quantifying Energy and Delay

The study of logical correctness and power requirements in Ch. 4.2 forms the basis for an expanded study that also considers energy and delay. In order to perform this study, first presented are the general expressions for quantifying the energy and delay of NML circuits as a function of the clock scheme used, as well as other necessary design parameters. These expressions – in addition to the study of clock schemes in Ch. 5.2 – will allow a more detailed analysis of the trends associated with magnet scaling, spacing scaling, and yoke
variation introduced in Ch. 4.3 as well as to make performance projections to compare to CMOS equivalents.

5.1.1 4-, 3-, 2-phase Clock Schemes

In an NML circuit, a multi-phase clocking scheme (such as those briefly examined in Ch. 3.3) can help to ensure that (a) signals can be “restored” such that a magnet can produce enough flux to properly bias the neighboring device that it is supposed to drive, and (b) back propagation is avoided. The multi-phase clocking scheme may use a different number of phases to achieve the aforementioned end goals. Fig. 5.1 illustrates how two adjacent clock wires could be excited under three different clocking schemes: 4-phase, 3-phase and 2-phase. For the 4-phase clocking scheme (Fig. 5.1a), current is ramped up in a clock wire, held at a maximum, ramped down, and then removed. The signal in each subsequent clock wire follows the same pattern but lags by 90 degrees. With a 3-phase clock, rather than ramping the current down from a peak value, the clock wire is abruptly shut off (see Fig. 5.1b), and the phase shift between two adjacent wires is 120 degrees. In a 2-phase clock, signals again are abruptly shut off but adjacent clock wires are excited out of phase (for a 180 degree phase difference).

Each clocking approach has inherent benefits and drawbacks. Some can take better advantage of NML’s inherent pipelining\(^1\) offering higher throughput, while others offer lower energy. The complexity of clock drive circuitry and/or a clock distribution network is another important factor. In the context of this study, of particular interest is how the different clocking schemes impact the functionality, performance and energy consumption of NML circuits. Specific quantifiable metrics will be discussed further in Ch. 5.2.

\(^1\)An NML circuit is “inherently pipelined” in the sense that once data propagates completely through a clock group (i.e., an ensemble of magnets controlled by one clock wire), new input signals can begin to traverse that same group. Because magnets are non-volatile and retain their binary state when no clock field is applied, data can be latched where it is needed.
Figure 5.1. Excitation patterns for (a) 4-phase clock, (b) 3-phase clock, and (c) 2-phase clock. $t_\phi$ is the phase time, i.e., the time duration of each phase. The red (solid) and blue (dashed) curves show the excitation pattern for two adjacent clock wires. The devices patterned on the clock wires represented by the solid red curve will drive the devices on the wires represented by the dashed blue curve. [25].

5.1.2 Energy and Delay Computation

Here general expressions are derived to quantify the performance and energy associated with NML circuits given the multi-phase clocking schemes discussed in Ch. 5.1.1. It is assumed that an ensemble of magnets is controlled by $N$ clock wires (clad with permalloy on three sides, as described in Ch. 2) and the number of magnets along the critical path on each clock wire (i.e., within each clock zone) is $M$.

In terms of NML circuit performance, both latency and throughput are of interest. Note again that the operating principle of NML circuits provides free pipelining. The latency of an NML circuit (i.e., the time required for a signal to propagate from a circuit input to output) can be computed by Eq. 5.1:

$$t_{lat} = \phi \times t_\phi + (N - 1) \times t_\phi = (N + \phi - 1) \times t_\phi$$

(5.1)

where $t_\phi$ is the time duration of a phase and $\phi(=2,3,4)$ represents the different multi-phase clocking schemes discussed above. If the dataflow patterns associated with an application of interest can leverage an NML circuit’s inherent pipelining, a new result can be produced every time the last wire in the circuit cycles through all $\phi$ phases. If this time is denoted as
as \( t_P \), then

\[ t_P = \phi \times t_\phi. \]  \hfill (5.2)

The inverse of \( t_P \) is then equal to the throughput of the circuit. The above expressions clearly show that the performance of NML circuits is critically dependent on the parameter \( t_\phi \). The value of \( t_\phi \) is impacted by \( M \) (the number of magnets controlled by each clock wire), the physical parameters associated with the clock wire, as well as the choices of the multi-phase clocking scheme. Simulation results to be presented in later sections will quantitatively demonstrate the effect of these parameters on NML circuit performance.

In terms of energy, there are three consumers in an NML circuit: the nanomagnets, the clock wires, and the clock drive circuitry. That is,

\[ E_{\text{total}} = E_{\text{mag}} + E_{\text{wire}} + E_{\text{cir}}. \]  \hfill (5.3)

Consider first \( E_{\text{wire}} \). The energy associated with the clad-wire clock can be attributed to the inherent wire resistance \( R \) and parasitics. \( R \) can be computed as:

\[ R = \frac{\rho l}{w h} = \frac{\rho l}{w^2 \times AR} \]  \hfill (5.4)

where the clock wire geometry is captured by \( w, h, \) and \( l, \rho \) is the resistivity of the copper, and \( AR \) is the aspect ratio of the wire. To reduce \( R \), thicker wires (i.e., wires with a large \( h, \) or \( AR \) value) are desired (an active fabrication target in [3]). The energy consumed by a clock wire in one cycle can thus be obtained as

\[ E_{1-\text{wire}} = \int_0^T P(t) \, dt = \int_0^{t_P} f^2(t) \frac{\rho l}{w^2 \times AR} \, dt \]  \hfill (5.5)

where \( T \) is the period of the clock.

In a given NML circuit that requires \( N \) clock wires (i.e., a circuit that has \( N \) clock zones), the energy consumed by all the wires for evaluating a single output in a non-
pipelined fashion (i.e., with each wire being excited only once) can be calculated as

$$E_{\text{wire-}np} = \int_{0}^{t_{\text{lat}}} I^2(t) \frac{\rho l}{w^2 \times AR} dt = \int_{0}^{\phi t_{\phi}} I^2(t) \frac{\rho l}{w^2 \times AR} dt + \int_{t_{\phi}}^{(\phi+1)\times t_{\phi}} I^2(t) \frac{\rho l}{w^2 \times AR} dt + \cdots + \int_{(N-1)\times t_{\phi}}^{(N+\phi-1)\times t_{\phi}} I^2(t) \frac{\rho l}{w^2 \times AR} dt$$

$$= N \times E_{1-\text{wire}}.$$  \hspace{1cm} (5.6)

If the circuit is used in a pipelined fashion, all $N$ wires are excited periodically with period $t_p$. If one ignores the pipeline filling and flushing overhead, $N$ new outputs are produced every $N \times t_p$ time duration. Therefore, the energy consumed by all clock wires for producing one output can be obtained as follows:

$$E_{\text{wire-p}} = N \times (N \times E_{1-\text{wire}})/N = N \times E_{1-\text{wire}}.$$  \hspace{1cm} (5.7)

Though the energy consumed by wires for producing one output is the same for both the pipelined and non-pipelined operating modes, the pipelined operating mode has much higher throughput and hence leads to much more desirable energy delay products (EDP).

Similar to CMOS circuits, the energy consumed by clock wires is strongly dependent on the current $I(t)$ passing through the wire. However, different from CMOS circuits, the required current $I(t)$ in NML circuits is ultimately determined by the magnitude of the magnetic field required to facilitate switching in an ensemble of magnets, which in turn depends on the number, size and spacing of the magnets on each clock wire, the multi-phase clock scheme used, and the actual clock phase time $t_\phi$. Due to the complex relationships among the parameters involved, it is more accurate to leverage micromagnetic simulation to obtain the magnitude of the magnetic field. Additional finite element simulations can then be used to determine the necessary current $I(t)$. Then, wire energy can be calculated following the above expressions. This is the strategy followed.

As shown in Eq. (5.3), the energy consumed by the transistor-based clock drive circuitry
and the switching energy associated with the magnets themselves should be considered for completeness. A voltage-controlled current source is needed to drive the clock wires. A CMOS dual complementary pulse current source with rise and fall time control was designed ([4]). Unlike traditional pulse current sources, for the purposes of this study, the requirement for a specific peak current output will not be that strict. It should be sufficient to have a peak current larger than the threshold current required for the minimum magnitude of the required clocking field. A single current source should be able to drive a large number of magnets/wires. To capture the impact of this circuitry on system-level energy, Eq. 5.7 is augmented to consider drive circuitry overhead. From initial analysis of the clock wire drive circuitry, it is estimated that 1/3 of the voltage will drop across the transistor and 2/3 will drop across its load (e.g., the clock wires). A conservative overhead of 50% is used in these calculations.

Finally, per [19], the impact on system-level energy from magnet switching events should be very low (e.g., on the order of 100-200 kT per magnet) – and will be dwarfed by the drive circuitry. As such, a foremost design goal is to minimize the energy of the required drive circuitry in order to realize the significant energy savings that magnetic logic could provide.

5.2 Comparison of Clocking Schemes

As noted in Ch. 5.1.1, there are multiple ways to excite ensembles of clock wires. Here, how different clock schemes will affect the logical correctness, latency, and energy of NML circuits is considered. How the different clock schemes ensure uni-directional dataflow is also discussed. Lines of magnets are investigated, but the results are also applicable to larger and more complex NML circuits.
5.2.1 4-Phase Clock

Recall from Ch. 5.1.1 that, given a 4-phase clock scheme, clock wires that control the magnets in groups (i) and (ii) in Fig. 4.1a would be excited simultaneously. An advantage of this approach is that a magnet at the end of a given clock wire group should only see one strong driver, and potentially unwanted feedback (discussed in Ch. 2 and Ch. 4.1.1) can be eliminated.

Considered first is how data propagates through lines of $40 \times 60 \times 20 \text{nm}^3$, $50 \times 75 \times 25 \text{nm}^3$, and $60 \times 90 \times 30 \text{nm}^3$ magnets with 8, 10, and 12 nm spacings between devices, respectively. These initial simulations allow estimates to be made of the length of $t_\phi$ required to facilitate propagation using a 4-phase clock. However, in order to study the energy and delay of these circuits, $t_\phi$ needs to be determined explicitly via simulation. As such, further considered is how the state of magnets in AF-ordered lines of $40 \times 60 \times 20 \text{nm}^3$, $50 \times 75 \times 25 \text{nm}^3$, and $60 \times 90 \times 30 \text{nm}^3$ changed in time as a function of how the 4-phase clocking field was applied (i.e., the simulation analysis step cutoff is an explicit time, and is not based on the change in magnetization per unit time). In order to understand the time-dependent behavior of the NML lines, a stopping criterion is used that allows precise choice of the length of $t_\phi^3$. As a specific example, consider an inverter chain of $40 \times 60 \times 20 \text{nm}^3$ magnets. The non-uniform field distributions shown in Fig. 4.1c were again considered. Simulations indicate that a minimum field of 3,979 A/m and a $t_\phi$ of 3 ns are required to ensure a new input signal can move beyond the second group of nanomagnets. (The specific waveform of the 4-phase clock is given in Fig. 5.1a, where each phase takes 3 ns and the peak field strength is 3,979 A/m.) Additional simulations found that $t_\phi$ could be

---

2 The device aspect ratio was chosen to strike a balance between magnetization state stability and the required clock field: as aspect ratio increases, the devices become more stable, but a higher external field is also required to bias the devices in an ensemble along their hard axes (see Ch. 3.2). Furthermore, this aspect ratio is used in experiments.

3 Again, the non-time cutoff simulations provide the approximate minimum field strength, and an approximate value for $t_\phi$: further simulations are run with varying values for $t_\phi$ to more specifically determine the minimum phase time required for propagation.
reduced to 2.5 ns if the rise/fall time is reduced to 1.25 ns. These and other results will be leveraged further in Ch. 5.4 to consider how an NML-based design might perform at the system/functional unit level.

5.2.2 2-Phase Clock

As discussed in Chs. 2.4.5 and 4.1.1, magnets may be driven from two directions if a 2-phase clock is employed. However, a 2-phase clock may be employed and unwanted feedback eliminated if non-rounded-rectangle magnets are included in the AF-ordered lines (or other NML circuit structures). More specifically, a magnet’s shape can help direct flux more unidirectionally – i.e., toward the magnet it should drive, and away from the magnet it should not drive. For example, in the presence of the same external field, a trapezoid-shaped magnet generates less “backward/anti bias” than a rounded rectangle (see Fig. 5.2a, left). Moreover, this makes the device an even stronger driver of the device that it is supposed to control (see Fig. 5.2a, right). (Note the lack of symmetry between curves at left, right is due to the fact that the magnetization of the magnets may be slightly curved rather than perfectly ↑ or ↓.)

Initial simulations (with uniform and continuous fields) indicate that placing a trapezoid at the boundary of a clock wire group (see Fig. 5.2b) leads to successful signal propagation with a 2-phase clock. Lines consisting of only rounded rectangles, on the other hand, were not well-controlled. However, the field needed for the line with a trapezoid (19,000 A/m assuming a uniform, continuous clock) was 5X greater than that needed for a 4-phase, non-uniform clock. The reason for this increase in field strength can be explained by the shape of the trapezoid as compared to the rounded rectangles. Fig. 5.2c depicts a comparison between the external fields required to bias a trapezoid-shaped magnet and a rounded rectangle shaped magnet into a metastable state. In this case, the trapezoid has a higher aspect ratio than the rounded rectangle. The trapezoid initially requires larger fields before it begins to bias in the x-direction, and once the field reaches a threshold, the magnet
quickly “snaps” to the metastable state. The rounded rectangle follows a more gradual transition to the metastable state, but does not require as strong nulling fields. Therefore, when this high-AR trapezoid is placed at the boundary between groups, it will raise the required clock field for the entire system. However, if the aspect ratios for the two magnets are the same (see Fig. 5.2d), note that the point at which the trapezoid “snaps” to the metastable state is at a lower field (whereas the rounded rectangle follows a similar curve). As such, a trapezoid-based approach to 2-phase clocking could result in lower energy operation if the aspect ratio can be lowered.

Thus, at present, given an existing simulation setup, even though fewer wire switching events would be required, the larger current needed to facilitate switching would erase any energy gains.

5.2.3 3-Phase Clock

Upon closer examination of the simulation results in Ch. 5.2.1, it was apparent that all magnets in a given wire group were biased into a logically correct state at the end of the second clock phase. This led to investigation of whether a 3-phase clock might also lead to logically correct behavior with even lower energy, since clock wires will be excited for one less phase than in a 4-phase clock, e.g., in Fig. 5.1a-h. An added benefit would be higher throughput – as per Eq. 5.2 assuming $t_\phi$ remains the same, a new result will be produced more frequently. Using a line with $40 \times 60 \times 20$ nm$^3$ magnets as an example, switching is facilitated with the same fields required for a 4-phase clock, but with a 3 ns $t_\phi$.

Additional experiments using the 3-phase clock to further explore the design space are now presented. Although the 4-phase clock can function correctly with a 2.25 ns $t_\phi$ (while

\[ \text{Note that these simulations consist of 31 magnets (again, an input followed by three groups of ten magnets each) followed by a large block with a low aspect ratio. Due to the excitation pattern of the 3-phase clock (see Fig. 5.1b) the line is more susceptible to a signal being driven backward by the last magnets in the third group (which strongly retain their original state when unclocked). A block mitigates these effects, as it more closely models the nulling contribution provided by a fourth group of magnets. Note that if a block is employed with a 4-phase clock (in order to have a fair comparison) $t_\phi$ can be lowered to 2.25 ns (vs. 3 ns).} \]
Figure 5.2. (a) Comparison of forward biases (light curves) and backward biases (dark curves) for trapezoid shaped magnets (dashed curves) and rounded-rectangle (RR) shaped magnets (solid curves). Plotted here is the absolute value of the y-directed flux across a 40 nm line 10 nm away from a given magnet shape; (b) Trapezoids at the boundaries between clock groups to prevent back propagation; (c) Comparison of fields required to place a single trapezoid shaped magnet (dark curve) and rounded rectangle shaped magnet (light curve) into a metastable state. Note the trapezoid has a higher aspect ratio than the rounded-rectangle; (d) Comparison of fields required to place a single trapezoid shaped magnet (dark curve) and rounded rectangle shaped magnet (light curve) into a metastable state. In this case, both magnets have the same aspect ratio. [25].
the 3-phase clock requires a 3 ns $t_{\phi}$ and thus have comparable delay (for one output to be produced), the 3-phase clock wins in terms of energy due to its specific current waveform (see Figs. 5.1a,b). Therefore, the 3-phase clock is chosen for use in this work (and recommended for future work) because, from a system-level perspective, it is an improvement over the 4-phase clock in terms of EDP (i.e., 1.5 times lower). Note also that the 3-phase clock is an improvement over the 2-phase clock in terms of energy (as, at present, the 2-phase clock requires higher magnitude fields). Furthermore, like a 4-phase clock, the 3-phase clock satisfies the digital logic tenet requiring elimination of unwanted feedback (since adjacent clock wires are still excited simultaneously, resulting in a clock group seeing only one strong driver). The interplay between the spacing between nanomagnets, the width of the yoke sidewall, and the size of the magnets themselves when considering a 3-phase clock scheme will all be investigated. These experiments are discussed in detail in Ch. 5.3.

5.3 Impacts of Physical-level Design Parameters

In this section, the three design parameters first discussed in Ch. 4.3 are revisited – the size of the magnets, the spacing between the magnets, and the width of the yoke sidewall – to determine how they affect the minimum field strength and phase time required for logically-correct signal propagation in lines of NML devices. A 3-phase clock is assumed for all simulations in this section.

5.3.1 Methodology

The simulations presented in this section assume the same parameters as discussed in Ch. 5.2.3. That is, lines of 31 magnets terminated on the right with a block are used. As before, discrete timesteps are used as the stopping criterion in the simulation in order to precisely control the length of $t_{\phi}$. Again, success is determined by whether a signal can propagate completely through the second group of magnets, which indicates that fringing
fields from a magnet at the output of group one are sufficient to drive an equally sized group of $M$ devices. The effects of the following are investigated:

1. **Magnet size** – $40 \times 60 \times 20$ nm$^3$, $50 \times 75 \times 25$ nm$^3$, and $60 \times 90 \times 30$ nm$^3$ devices are considered.

2. **The spacing between magnets** – 8 nm, 10 nm, and 12 nm spacings are considered.

3. **The width of the ferromagnetic cladding surrounding the copper clock wires** – 30 nm, 50 nm and 100 nm widths are considered.

The study is not meant to be exhaustive, but rather to show how the realistic drive circuitry overhead and layouts in the NML layer can impact both system-level energy and delay. These simulation results will be most helpful as future experimental efforts that target NML drive circuitry evolve. Note also that the first two items impact the necessary width of the copper clock wires. Finally, though the results presented in Chs. 5.3.2-5.3.4 are presented in a consolidated tabular form, the simulations for obtaining these results were conducted methodically. That is, magnet size is fixed, for example, and simulations were run to determine the effects of changing spacing, and so on.

5.3.2 Magnet Spacing

The spacing between magnets is intuitively known to be an important design consideration; as magnets are moved farther apart, fringing field interactions are not as strong. (Conversely, as magnets are moved closer together, fringing field interactions increase.) Results in Ch. 4.2.1 scaled the spacing between magnets as the size of the magnets scaled – i.e., with $40 \times 60 \times 20$ nm$^3$ magnets the distance between devices was 8 nm, with $50 \times 75 \times 25$ nm$^3$ magnets the distance between devices was 10 nm, and with $60 \times 90 \times 30$ nm$^3$ magnets the spacing between devices was 12 nm. Here, first discussed are results from an expanded simulation space designed to address how different spacings could affect signal propagation and energy requirements given different sized magnets (a 30 nm yoke is assumed for all simulations unless otherwise noted).
Consider first an important trend in the context of applied clock fields: as the spacing between magnets decreases, coupling between magnets increases. As such, a net effect can be that a relatively weaker external field is needed to put the magnets into the metastable state needed for re-evaluation (a desirable outcome, as the energy overhead of the clock should then decrease). Of course, if the magnets are farther apart, there will be less x-directed flux between devices. This trend is apparent when comparing the external fields required to facilitate propagation in lines made from magnets with different footprint areas and spacing between devices.

In Table 5.1, representative results are presented assuming both uniform and non-uniform field distributions. Note that in Table 5.1, numbers in parentheses represent results using uniform field distributions. Dashes represent cases where a working solution could not be found. (Recall uniform fields as discussed in Ch. 4.1.1: they terminate precisely between the two magnets at the boundaries, and have the same magnitude across the entire NML clock group.) With a uniform field distribution, the magnitude of the external field needed for re-evaluation increases as the spacing between magnets increases. However, with a non-uniform field, changes in magnet spacing actually can make it impossible for a signal to successfully propagate through a line of magnets (for some design configurations).

For example, given a line of $40 \times 60 \times 20 \text{ nm}^3$ magnets and 12 nm spacing, signal propagation with non-uniform fields was not possible even with very high field strengths and phase times due to weakened coupling at increased spacing. However, as seen from Table 5.1, a signal can successfully propagate through this same NML line given uniform field

---

5 The $60 \times 90 \times 30 \text{ nm}^3$ lines spaced 8 nm apart are a special case discussed below.

6 Simulation efforts were halted when it was apparent that increasing the field strength was not helping the signal propagate further. For example, in many cases the field strength applied was so strong that it completely nulled the line and prevented propagation; using a higher field strength would simply not be beneficial. Similarly, efforts were halted when increasing the phase time did not result in an improvement in signal propagation.
TABLE 5.1
TRENDS ASSOCIATED WITH MAGNET SPACING AND SIZE WITH RESPECT TO THEIR EFFECTS ON THE MINIMUM REQUIRED FIELD STRENGTH AND PHASE TIME.

<table>
<thead>
<tr>
<th>Magnet Size (nm³)</th>
<th>Spacing (nm)</th>
<th>Min. Field (A/m)</th>
<th>Min. $t_\phi$ (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40 × 60 × 20</td>
<td>8</td>
<td>3979 (1989)</td>
<td>3 (3)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>19894 (11937)</td>
<td>2 (2)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>– (23837)</td>
<td>– (3)</td>
</tr>
<tr>
<td>50 × 75 × 25</td>
<td>8</td>
<td>– (3979)</td>
<td>– (2)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>15915 (11937)</td>
<td>6 (4)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>– (19894)</td>
<td>– (4)</td>
</tr>
<tr>
<td>60 × 90 × 30</td>
<td>8</td>
<td>– (−)</td>
<td>– (−)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>3979 (3979)</td>
<td>3 (3)</td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>– (11937)</td>
<td>– (3)</td>
</tr>
</tbody>
</table>
distributions. Therefore, while this NML layout is not fundamentally prohibitive, it must be carefully studied within the context of the realistic clock. Similar results are seen for the $50 \times 75 \times 25 \text{ nm}^3$ and $60 \times 90 \times 30 \text{ nm}^3$ magnets when the spacing is increased. However, for the latter, per Ch. [4.2.1] if each clock group is restricted to 8 magnets, successful propagation is possible. As discussed in Ch. [4.1.1] the output of a group of magnets in a wire ensemble can be a weaker driver than the input to that group; by reducing the number of magnets in each group, the output can drive all $M$ devices.

The above results suggest that one would seemingly want to make the spacing between magnets as close as possible. However, x-directed coupling between devices can also be too strong. For example, given a line of 31, $60 \times 90 \times 30 \text{ nm}^3$ magnets (again followed by a block) with 8 nm between devices, the ground state (i.e., the magnetization state with no clock fields applied) for this setup was a line of magnets all biased completely in the positive x-direction (i.e., ferromagnetic ordering). As the magnets are so close, their coupling in the x-direction is strong enough that they essentially mimic a long strip of magnetic material. The magnets cannot be biased into a logically correct state even when uniform fields are applied. Similarly, a signal could also not propagate through lines of $50 \times 75 \times 25 \text{ nm}^3$ magnets with 8 nm between devices when controlled by non-uniform clock fields. With this configuration, an applied field was either too weak (i.e., magnets never reached the required metastable state and the line was ultimately driven from both directions) or too strong (i.e., the hard axis clock field sufficiently suppressed a given device’s y-component of magnetization such that y-directed fringing fields from said device were insufficient to set the state of a neighbor magnet, as in Ch. [3.1.2]). However, as seen in Table [5.1] with a uniform field, data did propagate successfully through this ensemble.

Finally, the above discussion might suggest that spacing variations during fabrication need to be less than 2 nm in order for an NML circuit/line to operate correctly. The large increase in field requirement when increasing spacing (e.g., from 8 nm to 10 nm for the
40 × 60 × 20 nm³ magnet size) is due to decreased coupling between neighboring devices. As the spacing increases, the field exerted by adjacent magnets decreases, and a higher external clock field is required to put magnets in a metastable state to facilitate propagation. However, it should not be the case that manufacturing variation be kept under 2 nm. As mentioned earlier, the design space explored here was not meant to be exhaustive. Rather, the intent was to show that (a) the behavior of an NML circuit can change when controlled by the non-uniform field distributions expected from a realistic/manufacturable clock, (b) realistic field distributions can impact the energy overhead of the drive circuitry, and (c) that there is a “sweet spot” when considering magnet spacing (i.e., for a given magnet footprint, more closely spaced devices do not always lead to improved power performance). In other words, the spacings chosen for this study may not be representative of what would be used in fabrication, and this data should not be used to draw conclusions about fabrication parameters or the impacts of variation.

That said, if future work shows that NML circuits do in fact have strict manufacturing tolerances, note that magnet aspect ratio and/or shape can also be leveraged to mitigate the effects of non-uniform field distributions. Two examples are presented. First, a line was considered that consisted of 31, 60 × 105 × 30 nm³ magnets followed by a block, with 8 nm spacing between devices. (Recall that a similar line of 60 × 90 × 30 nm³ magnets stayed hard axis biased even if no clock field was applied.) With the higher aspect ratio devices, a signal can successfully propagate from the beginning to the end of the line with non-uniform 11,937 A/m fields and a $t_{phi}$ of 6 ns (assuming a 30 nm yoke width). Magnets of a different shape, e.g., the trapezoids discussed in Ch. 5.2.2 can also be leveraged to address some of the issues associated with magnet spacing. As indicated in Table 5.1, the 40 × 60 × 20 nm³ magnets with a 10 nm spacing require a large increase in field requirement as compared to those at an 8 nm spacing (due to reduced coupling between neighbors in the magnet ensemble). Simulations of 31 magnet lines consisting of all trapezoids with a 10 nm spacing indicate that signal propagation can be achieved with a non-uniform field
of 7,958 A/m, and a $t_{\phi}$ of 2 ns (again assuming a yoke width of 30 nm). Note that the footprint of the trapezoids was $40 \times 60 \text{ nm}^2$, with a thickness of 20 nm – the same parameters as the rounded rectangles. These techniques (among others) can be used to design NML circuits if manufacturing tolerances become an issue.

5.3.3 Yoke Width

Theoretically, as the width of a clock wire’s sidewall cladding (or yoke) is scaled down, it should be easier for a signal to propagate through a line of magnets. As the yoke gets smaller, fewer magnets will lie over the yoke and be subjected to the field minima seen in Fig. 4.1c. Ch. 4.3.1 suggested that if yoke width increases – e.g., from 50 nm to 100 nm – the external field strength required to control the exact same line of magnets could double. (Note that these results considered a $d\mathbf{m}/dt$ cutoff as discussed above.) An expanded simulation space (using a time cutoff) suggests that increasing the yoke width should not have such a drastic impact on energy requirements. These results are summarized in Table 5.2. Note that in Table 5.2, dashes represent cases where a working solution could not be found, and that data is presented for the 30 nm yoke, followed by the 50 nm yoke, and then the 100 nm yoke (separated by commas). For example, given a line of $40 \times 60 \times 20 \text{ nm}^3$ magnets with 8 nm between devices, if $t_{\phi}$ is increased from 3 ns to 4 ns, a signal propagates successfully from the beginning to the end of the line (albeit with a decrease in performance). However, if spacing is increased to 10 nm, a signal will only propagate successfully through the line if a 30 nm yoke is employed. If yoke width is increased to 50 nm, the signal can only propagate midway through the second clock group – even when considering higher field strengths and a longer $t_{\phi}$. If the width is increased to 100 nm, the signal will not even propagate into the second clock group. Because the magnets are spaced 10 nm apart, which results in weaker coupling, minima over the clock wire cladding make it more difficult to place the last device in a wire group into a metastable state and
TABLE 5.2

TRENDS ASSOCIATED WITH SCALING THE WIDTH OF THE YOKE BETWEEN COPPER CLOCK WIRES.

<table>
<thead>
<tr>
<th>Magnet Size (nm³)</th>
<th>Spacing (nm)</th>
<th>Min. Field (A/m)</th>
<th>Min. $t_\phi$ (ns)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$40 \times 60 \times 20$</td>
<td>8</td>
<td>3979, 3979, 3979</td>
<td>3, 3, 4</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>19894, –, –</td>
<td>2, –, –</td>
</tr>
<tr>
<td>$50 \times 75 \times 25$</td>
<td>10</td>
<td>15915, –, –</td>
<td>6, –, –</td>
</tr>
<tr>
<td>$60 \times 90 \times 30$</td>
<td>10</td>
<td>3979, 3979, 3979</td>
<td>3, 3, 3</td>
</tr>
</tbody>
</table>

the system is more susceptible to the back propagation effects discussed in Ch. 2. Again, similar results are seen when considering lines of $50 \times 75 \times 25$ nm³ magnets.

Finally, for the $60 \times 90 \times 30$ nm³ lines, the required field strength and phase time is the same for all yoke widths – in fact, the field and phase time is the same when employing a uniform field distribution. This suggests that the larger magnets are less susceptible to the effects of the field discontinuities. Field distributions are essentially the same over the cladding for all magnet sizes; as such, larger magnets will have a smaller percentage of their volume resting in the region where the field strength drops, and will see less of an adverse affect. Therefore, as magnets are scaled further, placement over the cladding could have a more significant impact.

5.3.4 Magnet Dimensions

Data in Table 5.1 seemingly suggests that as magnet size increases, the field magnitude required to facilitate switching can decrease. Namely, magnets with a higher volume are capable of providing more x-directed flux to their neighbors to facilitate the switching/re-evaluation process (such that a lower magnitude external field would be required). However, scaling magnet sizes will also allow clock wire widths to scale. For a given clock wire
geometry, one could generate the magnetic field needed to control an NML-based circuit with less current. Per Ch. 2, this should further minimize the distance that magnetic flux travels in air, increasing the magnitude of the generated field.

In terms of throughput, theoretically, one would seemingly want to create the deepest pipeline possible – e.g., every magnet would be clocked individually. If one wire is divided into N wires, there should be no impact on power as, while there will now be N wires, with N times the resistance, the current in each wire will be reduced by $N^2$ (e.g., as $I = J \times \frac{w}{N} \times t$). However (in addition to obvious fabrication constraints), this trend is also limited as fields from the magnets themselves assist in the nulling/switching process. As such, smaller groups of magnets could require higher fields to switch. However, in the context of voltage controlled clocking via multiferroics (discussed in Ch. 2.2), every magnet is contacted and clocked individually. Though this represents a significant fabrication challenge, it could be a future path to lower energy operation and the deepest inherent pipelines possible.

5.3.5 Summary

From the above discussion, one can make some analogies to CMOS. In CMOS systems, increasing the supply voltage $V$ increases device power and decreases device switching times. In NML-based systems, increasing the magnitude of the applied field also leads to an increase in system-level power and can decrease switching times. However, the external field can ultimately impede switching as it can suppress a given device’s fringing fields which set the state of a neighbor. Moreover, performance gains from higher fields are often outweighed by higher energy clocks. As such, for the energy/performance projections in Ch. 5.4.1 efforts were made to first minimize this energy and then the associated delay was accepted.
5.4 Benchmarking

The study of lines of magnets presented in Ch. 5.2 and Ch. 5.3 is valuable for projecting the energy and performance of NML circuits at the circuit and architecture level. Said projections are presented in this section.

5.4.1 Circuit-Level Benchmarks

To investigate how an NML-based system might perform at the application level, as in [59], a 32-bit ripple carry adder (RCA) is used as a benchmark. Performance and energy projections will be based on (a) the energy and delay associated with different inverter chains following the methodology discussed in Ch. 5.2 and (b) a potential majority-gate based adder schematic shown in Fig. 5.3. This is not to say that the majority gates, wire crossings, and vertical IC designs needed for the adder have not been simulated. In addition to the experimental efforts reviewed in Ch. 2, vertical IC (Ch. 3.4) and in-plane crossings (Ch. 3.5) have been considered via OOMMF simulation. The simulation results presented in Chs. 3.4 and 3.5 show that the structures should be controllable with magnetic fields similar in magnitude to those needed for AF-ordered lines. Thus, though the AF-ordered lines are essentially a “best case” circuit in terms of energy and delay, parallels can be made to more complex circuits (though increased critical path lengths are taken into account).

However, assuming a 3-phase clock, a simulation of just a 30 magnet line requires around 1.5 days of simulation time on a high performance cluster. A similarly clocked, magnetic RCA design would be at least 10X larger and require at least 2 weeks of simulation time to test a single input combination. As such, in this study performance is derived from the critical path delay associated with, and the field magnitude to control AF-ordered

---

7The benchmarks presented here represent an earlier effort than those in, for example, [43]. The more recent results in [43] take into account inter-wire capacitance, clock wire routing schemes, etc. in addition to what is considered in this work. Therefore, the results in this work tend to be more optimistic, but are still important as they a) laid the groundwork for future benchmarking efforts, and b) show NML’s potential performance gains over CMOS (albeit under somewhat ideal circumstances). Finally, note that the benchmarking efforts are continually evolving and becoming more refined as the technology matures.
Figure 5.3. NML majority gate adder schematic. Note that the vertical interconnect, wire crossing (inset at left), and majority gates (M1, M2, M3) have all been simulated and/or experimentally demonstrated. [25].
lines, as they can be well-evaluated in a reasonable time frame.

The system-level design goals discussed in Ch. 5.2 (minimize energy first, then consider delay) are analogous to the design goals for subthreshold CMOS. As magnet switching frequencies will be on the order of 10 GHz at best [13, 36] comparing the performance of an NML-based circuit to a subthreshold equivalent makes sense (the technology should be well-suited for similar applications). Furthermore, due to this relatively low upper bound on switching frequency, NML is unlikely to fill the same application spaces as high-performance CMOS where delay is an important benchmark. In other words, the focus is on comparisons to CMOS technology that emphasize lowering energy over delay, and that map to low-power application spaces. As such, the data in Table 5.3—the energies and delays for subthreshold CMOS adders at the 65, 70, and 90 nm technology nodes—will be used as a basis for comparison (a comparison to 22 nm CMOS is made in Ch. 5.6.4).

To estimate the performance of an NML-based adder, note that the 1-bit adder design in Fig. 5.3 is 10 magnets wide. As such, data from 30 magnet inverter chain simulations (with 10 magnets per group) will be leveraged to determine the current required to produce a given clocking field. Data for $t_\phi$ can also be leveraged to consider expected latencies and the time that each clock line will be excited. $t_\phi$ is augmented to take into account that the critical path through a majority-gate based adder will be larger than that of an inverter chain. As mentioned in Ch. 5.1, the drive circuitry is conservatively estimated to increase system-level energy by 50%. The number of wires (N) was assumed to be 32 (as the width of the adder schematic in Fig. 5.3 is approximately 10 magnets), and each had an aspect ratio ($h/w$) of 3. Magnet energy was considered as in [19].

Current requirement is estimated using Ansoft’s Maxwell 2D as described in Section 4.1.3 i.e., it is known what current is required to generate a desired field for a given clock wire geometry. Note that the currents reported in [3, 2] are not representative of the currents that are ultimately anticipated for NML circuits. The switching experiments reported in
TABLE 5.3

PERFORMANCE PROJECTIONS FOR 32-BIT SUBTHRESHOLD CMOS ADD.

<table>
<thead>
<tr>
<th>Node (nm)</th>
<th>Supply mV</th>
<th>Energy$_{NP}$ (pJ)</th>
<th>$t_NP$ (ns)</th>
<th>EDP$_{NP}$ (pJ ns)</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>65</td>
<td>200</td>
<td>0.095</td>
<td>2953</td>
<td>281</td>
<td>[58]</td>
</tr>
<tr>
<td>70</td>
<td>184</td>
<td>1.48</td>
<td>5046</td>
<td>7468</td>
<td>[10]</td>
</tr>
<tr>
<td>90</td>
<td>200</td>
<td>0.6</td>
<td>6000</td>
<td>3600</td>
<td>[12]</td>
</tr>
<tr>
<td>70</td>
<td>293</td>
<td>5.0</td>
<td>488</td>
<td>2440</td>
<td>[10]</td>
</tr>
<tr>
<td>90</td>
<td>400</td>
<td>1</td>
<td>100</td>
<td>100</td>
<td>[12]</td>
</tr>
<tr>
<td>22</td>
<td>300</td>
<td>0.015</td>
<td>30</td>
<td>0.45</td>
<td>[66]</td>
</tr>
</tbody>
</table>

those papers required a very high clock field (and thus current). For the devices considered in [3], the patterned magnets were spaced very far apart, did not benefit from coupling between devices, and are switched directly along their easy axes. Note that the magnets wereclocked along their easy axes for the purposes of the switching demonstration. The lines and gates considered in [2] were also spaced very far apart, and a high current was required to set the state of the input (a magnet whose easy axis was in the direction of the applied magnetic field). Simulation has shown that ensembles of magnets at lower spacings can be switched with much lower fields (again, due to the increased nulling contribution of neighboring magnets), thus requiring much lower current. For example, as reported in Table 5.1, a field of 3,979 A/m is sufficient to switch lines of 31 $40 \times 60 \times 20$ nm$^3$ magnets spaced 8 nm apart (as used in these benchmarks). This field can be generated with a current of 2.29 mA – much less than the currents reported in [3, 2].

Performance projections for $40 \times 60 \times 20$ nm$^3$ magnet lines are presented in Table 5.4. Data for devices with these feature sizes is presented as magnet sizes are on par with
the minimum feature sizes of 65/90 nm CMOS circuits. When magnets are surrounded by air/oxide, the expected energy for a 32-bit add is slightly better than a subthreshold CMOS equivalent. Moreover, the delay for an add that does not take advantage of inherent pipelining is about 25X better than subthreshold CMOS with sub-200 mV supplies. If an application can take advantage of inherent pipelining, delay will be approximately 200X better than CMOS equivalents. Pipelined CMOS was not considered as latching overhead overwhelms any performance gains. Data points assuming slightly higher supply voltages are also included in Table 5.3 (latencies improve more rapidly than energy gains decay). Even with a higher supply V, an NML-based circuit’s EDP should be approximately 9X and 96X better assuming a non-pipelined and pipelined implementation, respectively. Finally, note the last rows in each of Tables 5.3 and 5.4 will be discussed in Ch. 5.6.4.

5.5 Path to Lower Power: High Permeability Materials

It is possible to achieve lower-power operation by changing the material properties in the gaps between nanomagnets. All simulations have assumed that nanomagnets are surrounded by air (with the permeability of free space $\mu_0 = 4\pi \times 10^{-7}$ H/m). It could be possible to increase the ratio of flux density to magnetic field strength ($\mu = \frac{B}{H}$) by surrounding magnets with a different material to increase absolute permeability. While this design approach will require careful study to ensure that a given material does not affect the magnetic/computational state of the devices, candidate materials do exist. For example, Freescale has considered dielectrics for MRAM parts with embedded magnetic nanoparticles to increase a word/bit line’s field strength without increasing current [72, 73]. Proposed materials could increase the absolute permeability ($\mu = \mu_0 \times \mu_r$) as $\mu_r$ could range from anywhere between 2-to-30. Moreover, the fact that particle sizes can be below the superparamagnetic limit should help ensure that magnetic state is not unduly influenced. Potential energy savings from enhanced permeability materials are summarized in
Table 5.4 (see 3rd column) and could be significant. EDPs could be more than 1000X better than the best CMOS projection in Table 5.3. More recently, high permeability films have been considered in more depth [47, 48, 49, 50, 51]. Notably, experiment shows that a $\mu_r$ of approximately 4.5 can be achieved, and NML devices have successfully been fabricated with EPDs [48], resulting in decreased switching field requirements.

### 5.6 Path to Lower Power: New Magnetic Materials

Lower energy operation might also evolve from changing the material of the magnets. [13] has proposed magnets with a magnetocrystalline biaxial anisotropy where each magnet will have a shape-defined uniaxial anisotropy, as well as a biaxial anisotropy term to help improve hard axis stability. As [13] does not consider clocking, the added biaxial anisotropy helps to keep the ensemble of magnets in a metastable state and mitigates the effects of fabrication variations, magnetic “noise”, etc. This should increase the likelihood that each magnet will change only in response to its predecessor. That said, the devices in [13] should be controllable by local clocking mechanisms and could lead to lower power operation as they could be easier to null and require a weaker clock field to prevent devices...
from settling to an incorrect state before being driven by the correct output [60].

5.6.1 Initial Work

A clocked line of $15 \times 60 \times 90 \times 5$ nm$^3$ cobalt magnets was studied. Magnets were separated by 20 nm. For simplicity, a uniform clock field was assumed. The line of 15 magnets was separated into two groups, each with 7 magnets, and the first magnet served as the input. A local field of approximately 15,000 A/m was required to facilitate a logically correct re-evaluation with a new input. For a line of permalloy magnets (with the same size, shape, and spacings), an external field of about 25,000 A/m was needed to facilitate re-evaluation. As devices with a biaxial anisotropy have not been experimentally demonstrated (and seemingly need to be grown epitaxially), they have not been included in the benchmarking study. However, clocked biaxial devices could represent an alternative path to lower power, which will now be considered in more detail.

5.6.2 Related Work

Magnets with biaxial anisotropy were first proposed in [13] in the context of soliton switching. In this switching regime, all magnets are biased along their hard axes by a global clocking field, inputs are set, and when the field is removed the input signal propagates in order through the system (like a magnetic soliton). One of the main challenges in this approach is to determine whether magnets can remain biased along their hard axes for a long enough time after the field is removed such that they do not switch before the signal has reached them. Magnets with biaxial anisotropy were proposed in [13] in order to alter the energy landscape of the nanomagnets to have a local energy minimum at $0^\circ$ (as in Fig. 5.4). Thus, the magnets could remain biased along their hard axes for a time defined by the energy barrier between the local minimum and an $\uparrow/\downarrow$ polarization. This contrasts with the approach considered throughout this work where relatively small groups of magnets are controlled by the line clocks described in Ch. 2.2. In this approach, the magnets (i)
switch in the presence of the applied field, and (ii) do not need to remain biased along their hard axes for long periods of time. In the context of (i), this can lead to lower energy as magnets do not necessarily have to be completely hard-axis biased to enable propagation (whereas in the other approach, if this were the case they would settle to a strong ↑ or ↓ when the field is removed). However, magnets with biaxial anisotropy could be leveraged in the line clocking approach and possibly lead to lower energy, as preliminary simulations demonstrated in Ch. [5.6.1]. In the context of line clocks, the clock scheme could be such that the magnets switch in the presence or absence of an applied field. Thus, described next are initial efforts to switch devices with biaxial anisotropy in the presence of an applied field, followed by an improved approach where switching occurs in the absence of the applied field, and a discussion of tradeoffs between the approaches.

5.6.3 Enabling Signal Propagation

Efforts to enable signal propagation in lines of magnets with biaxial anisotropy using the line clock approach are now discussed. Consider first a line of 16 60 × 90 × 5 nm³ magnets spaced 20 nm apart. The first magnet acts as input to the line, which is broken into three clock groups and terminated with a block of magnetic material (as in Ch. [5.2.3] in order to mimic a fourth clock group). The uniform/ideal 3-phase clock depicted in Fig. [5.1] is employed with $t_\phi = 3$ ns. The magnetic material used is cobalt with a biaxial anisotropy term (K1) of 30,000 J/m³. The simulation was carried out at a temperature of 0 K. This configuration is modeled after that in Ch. [5.3] though thickness is reduced from 30 nm to match the simulation in Ch. [5.6.1] and in [13].

5.6.3.1 Switching in the Presence of an Applied Field

Simulation at varying field strengths shows that the minimum required field to clock this design is 25 mT (a granularity of 5 mT was used, i.e., a 20 mT field was insufficient to move the signal through the line). Fig. [5.5] contains snapshots from the simulation
Figure 5.4. Energy barrier for nanomagnet with biaxial anisotropy for various values of $K_1$. Data is captured using OOMMF. [26].
to demonstrate the effects of clocking this line (note the third clock group is omitted). References to the phases (first, second, etc.) are in reference to Fig. 5.1b from left to right, which is duplicated in each of Fig. 5.5a-c and highlighted for convenience. Before discussing this simulation further, a naming convention for each phase is introduced to assist the discussion (and that in Ch. 6). Consider Fig. 5.6 showing six clock groups of a long pipelined NML line with a 3-phase clock (which could thus represent two bits of data). Each clock group cycles through the three phases in Fig. 5.1b: (i) the first phase in Fig. 5.1b is denoted as the “buffer” phase where the clock is high to prevent back propagation into the adjacent clock group; (ii) the second phase in Fig. 5.1b is denoted as the “switching” phase where the group switches in the presence of the applied field in accordance with the previous adjacent clock group; (iii) the third phase in Fig. 5.1b is denoted as the “driving” phase where the strongly AF-ordered, unclocked magnets drive their signal through the subsequent “switching” group.

First, Fig. 5.5a shows the state of the line 1.5 ns into the first phase, i.e., the first and last clock groups have a 25 mT external field applied (i.e., the first group is in the “buffer” state, the second is “driving” and the last is “switching”). Note how the magnets become strongly biased along their hard axes. Next, depicted in Fig. 5.5b is the state of the line at the last time step in the second phase, i.e., when the first group is in the “switching” state and the second is in the “buffer state” and thus the clock field is applied to both the first and second clock groups. Again in this snapshot note how all magnets are strongly biased along their hard axes. Finally, Fig. 5.5c shows the state of the line after 0.7 ns into the third phase, i.e., when the first group is in the “driving” state and all magnets have relaxed to a strong AF-ordering, and the second group is in the “switching” state. The signal completely propagates in this way through all three clock groups.
Figure 5.5. Snapshots of micromagnetic simulation for line of magnets with biaxial anisotropy ($K_1 = 30,000 \text{ J/m}^3$). The third clock group and block are omitted for readability. Insets highlight current clock phase for first (solid red curve) and second (dashed blue curve) clock groups. Field strength is 25 mT and the 3-phase clock from Fig. 5.1b is employed. (a) The first clock group is in the “buffer” state (clock field applied), the second is in the “driving” state (no field applied). (b) The first clock group is in the “switching” state (clock field applied), the second is in the “buffer” state (clock field applied). (c) The first clock group is in the “driving” state (no field applied), the second is in the “switching” state (clock field applied).

Figure 5.6. Bit definition in a pipelined, AF-line for a 3-phase clock. [26].
5.6.3.2 Switching in the Absence of an Applied Field

Though the setup in Ch. 5.6.3.1 enabled logically correct operation, an observation about the switching process can lead to a lower-energy clocking method. Consider again Fig. 5.5b: at this point (the end of the second phase), the first clock group should contain strongly AF-ordered magnets in accordance with the state of the input so it can successfully drive the signal through the second clock group. However, what one sees instead is the signal does not start to propagate through the first clock group until the clock field is removed (at which point it propagates quickly, as the hard-axis-biased magnets are readily tipped to an ↑ or ↓ state). This contrasts with previous simulations in this work (e.g., of supermalloy magnets) where the magnets retain some state until they are biased by their left neighbor. This difference can be explained by considering Fig. 5.7 which shows the effects of adding biaxial anisotropy to a 60 × 90 × 5 nm³ cobalt device in isolation. Using the analytical formulas in [4], the demagnetizing energy is plotted as a function of angle of magnetization when hard-axis-directed fields from 0 to 40 mT are applied to the device. The plot is augmented by adding markers at the minima on each curve, which indicate the expected angle of magnetization of the device for the given applied field. Note how as the field is increased to 40 mT, the device with biaxial anisotropy is completely hard-axis biased, but at the same field strength the device with no biaxial anisotropy retains its original magnetization (albeit weakly). This observation is supported by micromagnetic simulation where lines of cobalt magnets with K1 = 0 J/m³ require a clock field of 45 mT, significantly higher than the 25 mT required for magnets with K1 = 30,000 J/m³.

Though the signal correctly propagates, the second “switching” phase is essentially completely wasted energy (though the phase is still necessary to ensure unidirectional dataflow per Ch. 5.2). As an alternative, one could remove the clock field after the first phase, i.e., as in Fig. 5.8 the field in the switching phase is reduced to 0 mT to let the signal propagate during the “correct” phase, leading to reduced energy. This waveform could result in energy being cut in half assuming the same $t_\phi$ can be used. Note the field
Figure 5.7. Comparison of $60 \times 90 \times 5 \text{ nm}^3$ cobalt device with biaxial anisotropy ($K_1 = 30,000 \text{ J/m}^3$) to cobalt device with $K_1 = 0 \text{ J/m}^3$ for increasing hard-axis field strength. +/- markers denote minima on each curve.
in the first clock group remains high until the field from the second clock group reaches its maximum – this ensures no back propagation can exist in the line while the field ramps up in the second phase (a phenomenon observed in simulations where this overlap did not exist). As in previous simulations with a 3-phase clock, this ramp time is 0.5 ns. The energy consumption of this method is thus not 50% of the original, but rather approximately 58%. Future work could study the potential for a smaller overlap to come closer to maximum energy savings.

In order to demonstrate the effectiveness of this new waveform, snapshots from a new simulation are presented. All aspects of the simulation are identical, including field strength and \( t_\phi \), but the new waveform is used. Fig. 5.9a-c contains the results at the same timesteps as in the original simulation. In this case, the signal propagates through the first clock group during the switching phase as expected. The signal propagates completely through all three clock groups in this manner. The maximum field strength required (again, lower fields were attempted) is the same as with the original 3-phase clock, thus leading to energy savings of nearly 50%.

5.6.3.3 Varying K1 Value

Previous work in [26] suggested that lines of devices with biaxial anisotropy could switch in the presence of an applied field, leading to lower energy. Simulations in [26]

![Figure 5.8. 3-phase clock excitation pattern for use with biaxial devices controlled by a line clock.](image-url)
Figure 5.9. Snapshots of micromagnetic simulation for line of magnets with biaxial anisotropy ($K_1 = 30,000 \text{ J/m}^3$). The third clock group and block are omitted for readability. Insets highlight current clock phase for first (solid red curve) and second (dashed blue curve) clock groups. Field strength is 25 mT and the 3-phase clock from Fig. 5.8 is employed. (a) The first clock group is in the “buffer” state (clock field applied), the second is in the “driving” state (no field applied). (b) The first clock group is in the “switching” state (clock field only applied briefly to ensure unidirectional dataflow), the second is in the “buffer” state (clock field applied). (c) The first clock group is in the “driving” state (no field applied), the second is in the “switching” state (clock field only applied briefly to ensure unidirectional dataflow).
were carried out with $K_1 = 15,000 \, \text{J/m}^3$, and lines contained only a single clock group. This result is in contrast with the results in Ch. 5.6.3 where switching could not occur until the clock field was removed, suggesting there may be a tradeoff between $K_1$ and energy. The effects of changing the $K_1$ value is now briefly considered.

Previous work in [26] has shown that devices with $K_1 = 15,000 \, \text{J/m}^3$ could lead to operation with significantly lower field strengths. Fig. 5.10 shows why this might be the case: a device with $K_1 = 15,000 \, \text{J/m}^3$ becomes hard-axis biased when an external field of 35 mT is applied. Compare this to a device with $K_1 = 0 \, \text{J/m}^3$ which retains its magnetization, and a device with $K_1 = 30,000 \, \text{J/m}^3$ (as in Fig. 5.7) that becomes completely hard-axis biased at 40 mT. Thus, the devices with $K_1 = 15,000 \, \text{J/m}^3$ should operate with a lower clock field. This observation is supported by micromagnetic simulation. The simulations presented in Ch. 5.6.3 are repeated with $K_1 = 15,000 \, \text{J/m}^3$. The clock schemes in Fig. 5.1b and Fig. 5.8 are both considered. In both cases, the field strength required is 20 mT. Therefore, in this case, there is no benefit to switching in the presence of the applied field (unless in a field decrease of less than 5 mT, or in terms of lower delay, which was not considered in this study but could be considered in future work).

To reconcile this with the results in [26], recall that lines in that work only consisted of an input, 5 magnets, and a block. Per Ch. 4.1.1 at least three clock groups are required for realistic simulation of NML ensembles. Thus, a reduction in field strength seen in a simulation of a single group of magnets (plus a block) is understandably not necessarily indicative of performance of longer lines. Furthermore, Fig. 5.11 shows snapshots from micromagnetic simulation when the original 3-phase clock is used. Note in this case, the first 3 magnets switch in the presence of the applied field. As in the case where $K_1 = 30,000 \, \text{J/m}^3$ (in Fig. 5.5), the majority of switching occurs only after the clock field is removed. Thus, in this case, a lower value of $K_1$ could lead to lower-energy operation as compared to the case where $K_1 = 30,000 \, \text{J/m}^3$, but not due to the ability to switch in the presence of an applied field. Future work could consider a larger design space that includes effects of
temperature, minimum \( t_\phi \), and more fine-grained field strengths to explore this further.

5.6.3.4 Supermalloy

Finally, the effects of clocking a line of supermalloy magnets using the two clock schemes are noted. To compare to the devices with biaxial anisotropy, lines of \( 60 \times 90 \times 5 \) nm\(^3\) magnets are considered. Note that the required field strength when employing the 3-phase clock is 35 mT for both the original and the modified 3-phase clocks. However, these results may not be indicative of behavior one would observe in typical lines of supermalloy magnets, which are generally spaced closer together than is possible for cobalt (as the higher \( M_s \) of cobalt leads to strong coupling and completely hard-axis biased lines at such close spacings). For example, previous work in Ch. 5.3 showed that the minimum (uniform as in this study) 3-phase clock field needed to enable signal propagation in a line of \( 31 \times 40 \times 60 \times 20 \) nm\(^3\) magnets spaced 8 nm apart (with a terminating block) was 2.5 mT. However, consider what happens when this exact setup (31 \( 40 \times 60 \times 20 \) nm\(^3\) terminated by a block, and 8 nm spacing between devices) is simulated, changing only the way the clock is applied (i.e., employing the modified clock in Fig. 5.8). This simulation shows that logically correct operation can be achieved, but the required field strength increases to 5 mT. Therefore, energy is 2X lower when switching in the presence of an applied field (as in Fig. 5.1b) versus switching in the absence of an applied field (as in Fig. 5.8). Thus, for the smaller, more closely-spaced supermalloy magnets, energy could be lower when switching in the presence of an applied field. This suggests that switching in the presence of an applied field may be beneficial for certain configurations, depending on material, size, spacing, etc. This should be studied in future work in the context of a larger design space study.
Figure 5.10. Comparison of $60 \times 90 \times 5$ nm$^3$ cobalt device with biaxial anisotropy ($K_1 = 15,000$ J/m$^3$) to cobalt device with $K_1 = 0$ J/m$^3$ for increasing hard-axis field strength. +/- markers denote minima on each curve.
Figure 5.11. Snapshots of micromagnetic simulation for line of magnets with biaxial anisotropy ($K_1 = 15,000 \text{ J/m}^3$). The third clock group and block are omitted for readability. Insets highlight current clock phase for first (solid red curve) and second (dashed blue curve) clock groups. Field strength is 25 mT and the 3-phase clock from Fig. 5.1b is employed. (a) The first clock group is in the “buffer” state (clock field applied), the second is in the “driving” state (no field applied). (b) The first clock group is in the “switching” state (clock field applied), the second is in the “buffer” state (clock field applied). (c) The first clock group is in the “driving” state (no field applied), the second is in the “switching” state (clock field applied).
5.6.4 Summary

This section demonstrated how to clock lines of magnets with biaxial anisotropy using the line clock approach considered throughout this work. Notably, devices with biaxial anisotropy can operate with lower field strengths than devices without biaxial anisotropy. This is not only due to the fact that they are inherently easier to hard-axis bias, but also can take advantage of a modified, low-energy clock scheme. Ch. 6 will describe how the energy savings from introducing the biaxial anisotropy term are especially important in realizing high stability, low energy designs.

Finally, note the last columns in Tables 5.3 and 5.4. The CMOS and NML projections are based on updated analysis presented at the 2011 MIND Annual Review (and published in [66]). This analysis is based on the method developed in this work (note they are only reported here and not claimed as part of this work). The NML projections consider the following: a pipelined, 3-phase line clock; magnets with biaxial anisotropy; a 600 ps nulling time and 142 ps switching time per magnet; a 7 magnet critical path; a clock wire area that is 9 magnets tall by 4 magnets wide; a 2 mA current; and 16% clock driver overhead. The results show that even with the new, lower-EDP CMOS projections, an NML implementation could be approximately 128X better in terms of EDP. Thus, magnets with biaxial anisotropy could be a method to significantly decrease power consumption.

5.7 Conclusions

In this chapter, I showed how clocking theory can ensure uni-directional dataflow (Ch. 5.1-5.2) to satisfy the tenet for digital logic presented in Ch. 2.4.5. The clocking theory also allowed quantification of the energy and delay of NML inverter chains (Ch. 5.3). This work is crucial to NML as it allows for benchmarking comparisons that show NML to outperform low-power CMOS (Ch. 5.4.1). Additional paths to even lower power were considered in Ch. 5.5 and Ch. 5.6. Furthermore, the study enables the future development
of an NML shift register based on the well-studied inverter chains (in Ch. 6).
CHAPTER 6

NML SHIFT REGISTER

One of the main goals of my work is to explore NML at the architectural and application levels, (beginning with the study of adder benchmarking in Ch. 5.4.1). Throughout my work I have considered NML inverter chains, which have been proposed to be employed mainly as local interconnect. Another potential use for this simple circuit is as a shift register. The multi-phase clocks considered throughout this work enable inherently-pipelined operation (described in Ch. 5.1.1), which in turn enables shift-register operation in NML lines. When the proposed clad-wire clock mechanism considered in my work is employed, NML circuits are broken into groups controlled by each clock wire. The multi-phase clocks described in Ch. 5.1.1 control the excitation patterns of each wire. Thus, with the clad-wire approach, when an N-phase clock is employed it creates an inherent pipeline, and a single bit can be stored every N clock wires and act as a shift register. This idea is extensible to other clock implementations, but again I focus on the clad-wire approach as it has been physically demonstrated [3, 2], and can offer performance improvements compared to CMOS equivalents (as in Ch. 5.4).

An NML shift register can be designed to be volatile (state is lost when un-powered) or non-volatile (state is preserved when un-powered). A non-volatile shift register has many potential useful applications, including a general checkpointing application (discussed in Ch. 6.8.1) to save computational state (typically, for long-running processes) and for instant-on operation. A shift register is also necessary to implement some NML applications, including systolic convolution (described in [18]) or a systolic pattern matcher – in either case, non-volatility is optional but could protect against data loss from power
failure (thus saving computation time). Furthermore, non-volatility is one of the primary advantages that NML has over CMOS equivalents. Realization of non-volatile logic is an active target being pursued by, for example, the DARPA NV Logic program [22], as non-volatile logic in general can be a path to extending the performance improvements associated with Moore’s Law as CMOS scaling approaches fundamental limits. Therefore, a significant portion of this chapter considers how to enable non-volatile shift register operation in order to enable specific applications of interest (namely, the checkpointing application discussed in Ch. 6.8.1). As I will demonstrate in Ch. 6.7 there is a tradeoff between non-volatility and energy, and I describe how to study these tradeoffs to achieve the desired balance.

6.1 Initial Work

Early simulation efforts with NML shift registers considered a line of forty $40 \times 60 \times 20$ nm$^3$ supermalloy magnets spaced 8 nm apart and divided into groups of five magnets each. Recall that lines of $40 \times 60 \times 20$ nm$^3$ magnets with 8 nm between devices (i) allow for a reasonable comparison to 15 nm CMOS when considering minimum feature size, (ii) should be large enough not to be in the superparamagnetic state, (iii) could be controlled by a line clock with non-uniform field distributions (Ch. 4.1), and (iv) the currents required to generate sufficiently high clocking fields should allow for clock energy budgets to make NML competitive with low power CMOS equivalents (Ch. 5.4.1 [18]). Simulations were again carried out using OOMMF [27], and thermal noise was modeled with random fields, whose strength is a function of temperature, that are applied to each mesh point for each simulation time step. A uniform 3-phase 5 mT clock drove inputs through the line successfully – temperatures of 0, 300, 500, 600, and 700 K were considered. With a 3-phase clock, this shift register can hold three distinct bits, and the pattern of inputs was varied so every permutation of the three bits was tested. These simulations showed for the first time that an NML inverter chain can successfully store multiple bits and act as a shift register.
In addition to logically-correct clocked operation, methods are needed to ensure non-volatility in the shift registers for applications where it is desired. In Ch. 6.2 I describe why shift registers may not be non-volatile despite the fact that NML devices in isolation are expected to be. In Ch. 6.2.2 I demonstrate the impact of this volatility on the steady state stability of the NML shift registers. Next, in Ch. 6.3 - 6.4 I outline methods to improve stability, and a method to quantify the stability of NML ensembles (described in [43]). Using these methods, I describe a design space sweep I performed to identify stability (Ch. 6.5) and energy (Ch. 6.6) trends, as well as the tradeoffs between stability and energy (Ch. 6.7). This study can be used as a guide for future NML circuit design efforts. Finally, in Ch. 6.8.1 I describe how NML shift registers might be used in two checkpointing applications, but note that performance will ultimately be decided by the behavior of NML I/O structures (currently under development).

6.2 Volatility in NML Ensembles

In this section, I describe the origins of volatility in NML ensembles. Namely, after discussing related work, I show how bit conflicts in defect-free, pipelined circuits can result in a state-destroying random walk.

---

1To be explicitly clear about contributions: the work in this chapter was in large part a collaboration with Steve Kurtz, and resulted in a co-first-authored publication [26]. My contributions are: (a) initial simulations demonstrating the random walk in defect/error-free, pipelined ensembles, (b) all efforts to clock lines (shift registers) containing latch structures, and to identify energy requirements of the structures, (c) the modified clock waveforms required to facilitate dataflow through latch structures, (d) the design space study of stability and energy of shift registers presented in Ch. 6.5-6.7. Steve’s contributions to this work are: (a) a demonstration of how coupling between magnets in ensembles affects the energy barrier of magnets in the ensembles, (b) the three architectural approaches to improving stability summarized in Ch. 6.3 (c) the methodology (and corresponding code) to calculate relative stability of NML ensembles, (d) Steve also presented an initial design space sweep considering only stability that produced identical results. The different between (d) in my work and (d) in Steve’s work is thus the context (stability in Steve’s work versus stability and energy tradeoffs in my work).
6.2.1 Related Work

Non-volatility in NML has largely been quantified by considering the energy barrier (EB) of a single device in isolation. However, device-to-device coupling will also affect a circuit ensemble’s overall stability. To begin, consider work published in [14] – which suggests that NML devices could operate near thermal equilibrium without the assistance of externally applied magnetic fields. [14] initially considered micromagnetic simulations of $60 \times 120 \times 5$ nm$^3$ devices with a device-to-device spacing of 20 nm. This line was initialized such that all devices were initially in a $0^\circ$, metastable state, and was subjected to a hard axis magnetic field (37 mT) such that the EB between $\pm M_y$ magnetization states (i.e., $\uparrow$ and $\downarrow$ states) was less than the thermal energy $kT$. While all devices in the line relaxed to a magnetization state defined by an easy axis (i.e., $\uparrow$ or $\downarrow$), thermal noise caused some devices in the line to switch prematurely. The net result was a line with defects (where “defect” suggested improper signal ordering – e.g., $\uparrow\downarrow\uparrow\downarrow$ – rather than manufacturing variation). With continued application of the critical switching field, said defects began to randomly walk through the line (i.e., $\uparrow\downarrow\uparrow\downarrow$ might change to $\uparrow\uparrow\uparrow\downarrow$, etc.) until all mis-orderings were eventually annihilated. This phenomenon was repeated experimentally where temperature elevations alone induced bit flips (detected by photoemission electron microscopy). In essence, conflicts migrated, and in turn changed the state of surrounding devices.

6.2.2 Impact of Bit Conflicts in the Steady State

Part of an initial study of clocked operation of NML shift registers was to consider whether the circuits could maintain state at room temperature when no clock was applied. Clocked NML is usually inherently pipelined (e.g., see Ch. 5.1.1 [31], etc.) in order to improve throughput (as shown in Ch. 5.4.1 this helps to make NML competitive with CMOS at the application level). Recall that pipelining also enables shift-register operation. However, pipelining will lead to bit conflicts during error free operation (i.e., even with no
premature switching). Consider a line clock where an odd number of magnets span the width of each line. From Fig. 6.1a, if a group stores a binary 0, and an adjacent group also stores a binary 0, there cannot be complete AF-ordering in the ensemble. Similarly, if an even number of magnets spans each group, a binary 1 next to a binary 0 also results in a steady state conflict (Fig. 6.1b).

To see the potential impact of conflicts in pipelined circuits, considered again are micromagnetic simulations of a line of 40, $40 \times 60 \times 20$ nm$^3$ supermalloy magnets spaced 8 nm apart. As described above, simulations at various temperatures suggest that multiple bits can simultaneously move through the line as fields were applied to groups of 5 magnets to enable pipelined operation (recall that with a 3-phase clock, the line could store 3 bits of information as in Fig. 5.6). As in Ch. 5.4.1 a 5 mT clock field was sufficient for re-evaluation.

Now, consider the state of this line of magnets if adjacent magnet groups contain logic 0s. With this correct “architectural state,” bit conflicts occur. Lines with this conflict state were again considered via OOMMF simulation. No hard-axis fields were applied, and the ensemble was only subjected to the effects of thermal noise representative of a 300 K environment. In simulations with three different random seeds, random walks occurred. As an example, Fig. 6.2 shows a snapshot of the random walk. Here, after approximately 15 ns, the final state of the 40 magnet line that originally contained conflicts is a completely AF-ordered line, and information originally contained in the line is lost. Simulations of other devices/spacing (e.g., $60 \times 90 \times 30$ nm$^3$ magnets with 10 nm spacing) also exhibit state-destroying random walks.

Improvements to device and ensemble stability are necessary in order to achieve non-volatility – again, desirable to allow for interesting applications (checkpointing, instant-on operation, etc.) and to realize any potential energy savings compared to CMOS equivalents. Methods are needed to: 1) demonstrate that the pipelined NML shift registers can maintain the state of data in the circuit when power is removed, 2) show that computation can be

123
Conflicts occur if (a) an odd number of magnets per clock group or (b) an even number of magnets per clock group is used. [26].

In order to understand these methods, one must first understand how stability is measured. For a magnet in isolation, stability is typically presented as the energy difference between states (measured in kT). For example, consider the black curve in Fig. 5.4 – the stability of the $60 \times 90 \times 5 \text{ nm}^3$ magnet is defined by the energy difference between the $\uparrow$ ($90^\circ$) and $\downarrow$ ($-90^\circ$) states. The calculated stability depends on the temperature, and can...
be used to calculate an expected stability time for the state of the magnet. For example, a commonly desired value (e.g., in magnetic storage media) is around 40-60 kT at room temperature; 40 kT roughly corresponds to a stability time of 7.5 years. Increased kT results in exponential time increase, e.g., 45 kT corresponds to approximately 1100 years at 300 K. By Arrhenius-Neel theory of thermally activated magnetization reversal, the probability per unit time of reversal over the EB is $1/\tau$, where $\tau = t_0 e^{(EB/k_B T)}$ (where $t_0$ is assumed to be 1 ns per \[79\]).

There are straightforward methods to increase the stability of a nanomagnet. For example, changing material parameters can alter the energy landscape. Consider the green curve in Fig. 5.4 where a biaxial anisotropy term is added to the energy as first proposed in \[13\] (for description of the effects of biaxial anisotropy, see Ch. 5.6). Alternatively, changing the shape (e.g., increasing aspect ratio) of a magnet can similarly result in increased stability by raising the height of the EB in Fig. 5.4. In both cases, there will be an impact on the field required to null (i.e., place the magnet in a $0^\circ$ metastable state as required for re-evaluation) the nanomagnet, which can be expressed by the following equation from \[4\]:

$$H_{null} = \triangle N d M_s = (N_{dx} - N_{dy})M_s = (2t/\pi w - 2t/\pi h)M_s \quad (6.1)$$

Here, $N_{dx}$ and $N_{dy}$ are the demagnetizing factors in the x- and y-directions, respectively.
$w$, $h$ and $t$ are the dimensions of the magnet, and $M_s$ is the saturation magnetization of the magnet. The nulling field (i.e., the clock field required to bias a device along its hard axis) required thus depends on the dimensions of the nanomagnet, as well as its material properties. Thus, going forward, when explaining why design changes impact stability and required clock field, frequent reference will be made to the energy landscape of the magnets (as in Fig. 5.4 and Eq. 6.1).

These design levers are used to create high-stability structures to ensure stability at the architectural level, i.e., specifically to address the issue of instability in pipelined ensembles. By placing high-stability structures – henceforth called latches – in each clock group, architectural stability can be ensured. These high-stability latches must prevent the random walk described in Ch. 6.2.2 and store the state required to restart computation after some amount of time.

Before describing these structures in detail, recall the 3-phase clock described in Ch. 5.2, and that in the case of a 3-phase clock, each set of 3 clock groups contains a single bit in the pipelined shift register. Within each bit, there are three groups: one that is “driving”, one that is “switching”, and one that is in a “buffer” state (see Fig. 5.6). The driving group contains the state being shifted through the subsequent switching group. The buffer group remains clocked to ensure uni-directional dataflow as described in Ch. 5.2. These denotations are important when understanding how state is maintained with the latch structures.

6.3.1 All-Magnet Latch

The first approach, called the all-magnet latch (AML), is to make every magnet in the shift register a latch (see Fig. 6.3i). In other words, the stability of every magnet can be increased to a point where conflicts are confined to their original location (i.e., prevent the random walk) and maintain stability. Conflicts will be confined to the switching or buffer groups (depending on when power is lost), and only the first and last magnets in a driving group could have an adjacent conflict. However, the stability would be designed such that
the state of the magnets in the driving group should never be affected by these conflicts. One possible way to accomplish this is to increase the aspect ratio (AR) of all magnets in the line. To restart computation, each group resumes the role it had when power was lost (e.g., driving groups resume driving).

6.3.2 Single-Magnet, Three-Magnet Latches

Rather than increasing the stability of all magnets, one could instead place one or more high-stability latches in each clock group. For example, in the case of the single-magnet latch (SML), a single high-stability magnet is placed in each clock group, and the random walk is allowed to occur in all other, non-stable magnets (see Fig. 6.3ii). The SML is responsible for maintaining stable state even though it could potentially be in conflict with both its neighbor magnets (since conflicts can form in adjacent switching/buffer groups and walk through the low-stability magnets). Alternatively, one could place three adjacent high-stability magnets in each clock group, i.e., a three-magnet latch (TML). This would operate in the same manner as the SML, but now in the worst case each of the latches could only see a conflict from one side (see Fig. 6.3iii). That is, conflicts could randomly walk to either side of the TML, but each latch in the TML would have at least one high-stability neighbor. To restart computation, as with the AML, each group resumes the role it had when power was lost.

In the context of NML shift registers, these latches could be placed at the beginning, end, or somewhere in the middle of each clock group. Previous work using high-AR

![Figure 6.3. Latch methods: (i) AML (ii) SML (iii) TML.](image_url)
trapezoid-shaped magnets to enable 2-phase clocks (Ch. 5.2.2) suggested that the best place (in terms of field requirement and ease of achieving logically correct operation) for these devices was at the beginning of each group. In the case of the 3-phase clocks used in the context of latching, initial simulation work indicated again that they should be placed at the beginning of each clock group. This makes sense in the context of the discussion in Ch. 4.1.1 which showed how a magnet internal to a group is a weaker driver than the initial input to that group. Therefore if the latches have, for example, a higher aspect ratio than their neighbors, they may be harder to drive as (per Eq. 6.1) $N_{dy}$ will shrink due to increased $h$, thus resulting in a higher $\Delta N_d$, and consequently a higher biasing field (assuming the latches aren’t completely nulled, which is the case in this switching paradigm). However, internal drivers provide weaker biasing fields per Fig. 4.2, so latches placed internal to a clock group could require more energy to clock. By placing the latches at the beginning of a group, they see the strongest driver possible, which mitigates the impact of these issues.

6.4 Quantifying Stability

Now that the ways in which stability can be maintained at the architectural level have been discussed, a methodology is needed to quantify the stability of NML ensembles. Then, the three approaches can be compared, as well as other methods to improve stability (e.g., different materials). Ch. 6.4.1 is a high-level summary of the work in [43]. Thus, for details on theory, etc., please see [43]. Here, this methodology is used to study stability and energy in the context of NML shift registers and, specifically, to identify designs that provide high stability and low energy. In Ch. 6.4.2 two examples of how this methodology is used are presented, forming the basis for the design space study in Ch. 6.5.
6.4.1 Ensemble Stability

It is straightforward to quantify the stability of a nanomagnet in isolation using analytical methods (e.g., as in [4]), or using OOMMF. However, in order to design magnets to maintain stability despite architectural conflicts, one needs to consider the stability of ensembles of nanomagnets. Therefore, this work uses the analytical method described in [43] that estimates ensemble stability in the context of 3-magnet systems. Specifically, this method considers the effects of two neighbor magnets on a center “target” magnet. The effects of the fringing fields from the neighbors can be used to determine the stability of the target magnet. This is best illustrated with an example: if the target nanomagnet is supposed to remain in an \( \uparrow \) state, the analytical model calculates the EB between the desired \( \uparrow \) state and the opposite (\( \downarrow \)) state when the target is subject to fringing fields from the two neighbors. The EB (measured in kT) contribution of the target itself is also taken into account. This calculation is performed for all possible angles of the neighbors, resulting in an EB prediction for the target for each case.

As this method is based on the analytical formulas in [4] (including Eq. 6.1), there are some inherent limitations. First, there are some simplifying assumptions in the analytical formulas, e.g., that the magnets are single domain and rectangular. Also, this method alone cannot capture the dynamics of NML circuits that will exist in reality. Furthermore, it is not possible to verify results via micromagnetic simulation as it takes an inordinate amount of time to simulate even many ns of operation, whereas of interest are stabilities on the order of years. For these reasons, no claim is made that the analytical method can predict the absolute stability of an NML ensemble, but relative comparisons can be made between different designs to study trends and identify promising areas of future study (specifically, design levers that can offer high stability with low energy, and designs that should be studied via temperature-accelerated experiment).
6.4.2 Examples

The use of the analytical method from [43] is best described with examples of initial efforts to increase the stability of the NML shift registers. Also described is how these shift register designs are clocked, including field requirements and clock scheme used. An un-latched design, and the AML and SML are used as examples.

6.4.2.1 No Latching

As an example of how the analytical method is used, a 3-magnet system is considered that consists of $40 \times 60 \times 20 \text{ nm}^3$ supermalloy magnets with 8 nm spacing between each magnet (as in Ch. 6.2.2). Results are presented in Fig. 6.4, where each contour line represents ‘N’ in ‘N’ kT for a temperature of 300 K. This contour plot describes the predicted EB of the target magnet in the 3-magnet ensemble for every possible configuration of the two neighbors. Consider first quadrant I in Fig. 6.4, which represents the case where the two neighbors conflict with the state of the target. In this case, essentially no stability is predicted for the target (i.e., the target is expected to transition from $\uparrow$ to $\downarrow$). The case in quadrant I is only applicable in the context of the SML and TML, where the random walk could result in this worst case configuration.

Quadrants II and IV represent the case where only one neighbor conflicts with the state of the target (thus, this is applicable to the AML as well as the SML and TML). In this case, the stability of the target is such that it is likely to flip unless its neighbors remain at very high angles of magnetization (e.g., above approximately $\pm 60^\circ$ in Fig. 6.4 to remain above 40 kT). In reality, per the simulation in Ch. 6.2.2, the angles of the neighbors are such that the target flips and the random walk moves through the system. In fact, even at 0 K the angles of the neighbors tends to be approximately $\pm 45^\circ$, resulting in very low stability (see crosshair in Fig. 6.4). Finally, quadrant III represents the case where the neighbors are in agreement with, and reinforce the state of the target, and thus the stability of the target is very high. In terms of clocking, recall that when a 3-phase clock is employed, a 5 mT
Figure 6.4. Predicted EB (40 × 60 × 20 nm³ device, 8 nm spacing). [26].
field is required for this design.

In order to ensure non-volatility in the shift registers, one can, for example, employ an AML where the stability in quadrants II/IV is raised sufficiently high to eliminate the random walk (i.e., to confine conflicts to their original location). Alternatively, one can employ an SML or TML where the stability in quadrant I is raised high enough to withstand the worst case where the non-latch neighbor magnets oppose the state of the latch.

6.4.2.2 All-Magnet Latch

Next, consider an AML consisting of $40 \times 80 \times 20 \text{ nm}^3$ magnets with 8 nm spacing. Fig. 6.5 depicts the results of the analytical method for the EB of the target magnet in the 3-magnet system for every possible angle of its neighbors. Recall that the goal of the AML is to increase the stability of every magnet to confine conflicts to their original location. Thus, the worst case (quadrant I in Fig. 6.5) would never occur as the random walk would be eliminated. The challenge in designing the AML is to ensure that the ensemble stability is sufficiently increased to achieve this goal. In the case of the AML of $40 \times 80 \times 20 \text{ nm}^3$ magnets spaced 8 nm apart, the stability in quadrants II/IV is quite high and should prevent the random walk. Specifically, if the angle of magnetization of the neighbors remains above $\pm 45^\circ$ (see red crosshair in Fig. 6.5), the EB of the target is predicted to be over 100 kT (versus a few kT in the un-latched case, Fig. 6.4). This prediction is supported by micromagnetic simulation. The simulation is identical to that in Ch. 6.2.2 – i.e., bits are shifted through such that conflicts exist, the clock is removed and the state of the magnets in the line is observed – but with $40 \times 80 \times 20 \text{ nm}^3$ magnets. In this case, the conflicts remain confined to their original locations after 40 ns of simulation time (and thus the worst case, quadrant I, can be safely ignored). Recall the conflicts in the lines of $40 \times 60 \times 20 \text{ nm}^3$ magnets immediately began to randomly walk, and destroyed all saved state after only 15 ns. Finally, note that the angle of magnetization of the neighbors remains at an average of $\pm 86^\circ$ over the course of the relaxation simulation, suggesting
very high stability.

In order to re-evaluate these high-AR lines, a significantly higher clocking field is required than in the un-latched case. As the aspect ratio of a magnet goes up, it becomes more difficult to put into the metastable state required for re-evaluation (recall in Eq. 6.1, a higher AR results in higher $\Delta N_d$, causing $H_{null}$ to increase). Therefore, in this case the required clock field is 50 mT, resulting in a 100X energy increase. Note that other designs could offer high stability with significantly lower energy. For example, $40 \times 70 \times 20 \text{ nm}^3$ magnets require a clock field of only 25 mT, and are estimated to provide 500 kT stability if they remain magnetized at $\pm 83^\circ$.

### 6.4.2.3 Single-Magnet Latch

Finally, consider a line of $40 \times 60 \times 20 \text{ nm}^3$ magnets with a $40 \times 80 \times 20 \text{ nm}^3$ SML in each clock group. Fig. 6.6 depicts the results of the analytical method for the EB of the target magnet in the 3-magnet system (i.e., the $40 \times 80 \times 20 \text{ nm}^3$ magnet) for every possible angle of its neighbors (i.e., $40 \times 60 \times 20 \text{ nm}^3$ magnets). Recall in the case of the SML the goal is to design the latches to remain stable despite the fact that random walks in the remaining magnets could result in a worst case configuration (i.e., both neighbor magnets opposing the latch’s saved state, as in quadrant I in Fig. 6.6). Again, due to the limitations of the analytical method, one cannot guarantee these predicted EBs to be absolute. However, one can compare the relative stabilities of the SML and un-latched approach. As seen in Fig. 6.6, the $40 \times 80 \times 20 \text{ nm}^3$ latches provide significantly increased stability as compared to that in Fig. 6.4. Specifically, the stability in quadrant I of the former design is an average of 83X higher than the latter. Unless the angle of magnetization of the neighbors falls below $\pm 45^\circ$ (see red crosshair in Fig. 6.6), the latch should maintain its state. Therefore, the $40 \times 80 \times 20 \text{ nm}^3$ SML clearly provides a large increase in stability and may be able to enable a non-volatile shift register. Micromagnetic simulation again supports these observations: over 40 ns simulations, the state of the latches is not disturbed.
Figure 6.5. Predicted EB (40 × 80 × 20 nm$^3$ AML, 8 nm spacing). [26].
In fact, conflicts do not migrate at all over this time period due to the strong influence from the high-AR latches.

It should be noted that to re-evaluate these lines, simulations suggest that a unique clock scheme should be employed. The high-AR latches still require a significantly high field to allow for their re-evaluation. As such, the maximum clock field required is still 50 mT as in the case of the AML. However, once the latches are placed in the metastable state, the non-latch magnets can switch in the presence of a significantly lower field. This waveform is depicted in Fig. 6.7 where the first phase in this example operates at 50 mT, and the second phase operates at 5 mT. The result is a clock energy requirement that is 50X higher than the un-latched case – but half that of the energy requirement in the AML example. Furthermore, other designs could offer high stability with significantly lower energy. For example, \(40 \times 70 \times 20 \text{ nm}^3\) latches require a clock field of only 25 mT, and thus a 13X energy increase, while providing 19X higher stability in quadrant I compared to the un-latched case.

6.4.2.4 Summary

Clearly, stability is improved by employing the AML or SML, but one cannot guarantee an absolute stability due to the nature of the analytical method. In both cases, \(40 \times 80 \times 20 \text{ nm}^3\) or \(40 \times 70 \times 20 \text{ nm}^3\) latches could provide increased stability, but the question is which will provide enough stability to ensure shift register state is maintained. Experimental work is ultimately needed to validate these results and answer this question. For example, many lines of magnets (with varying aspect ratios) could be fabricated, initialized to a state with a conflict, and the state observed over time in order to calibrate the analytical model. For the purposes of this work in the context of shift registers, this analytical method can be used to compare relative stability between various designs to investigate tradeoffs between stability and energy to help guide future designs (especially with respect to future experiments). Next, Ch. 6.5 discusses efforts in this regard, and identifies promising designs to seed
Figure 6.6. Predicted EB (40 × 80 × 20 nm³ SML, 8 nm spacing). [26].
future experimental efforts.

6.5 Stability Trends

Using the analytical model, a design space sweep was first performed to determine how these approaches might impact stability. This discussion is closely coupled with the forthcoming Ch. 6.6 where energy trends are discussed for the TML, and Ch. 6.7 where metrics are described that identify the low energy/high stability configurations in this study.

The design parameters of interest (and space swept in parentheses) for the latch configurations are: spacing (8 nm to 12 nm in 2 nm increments), aspect ratio (1.5 to 2.25 in increments of 0.25), thickness (5 nm to 20 nm in 5 nm increments), magnet footprint ($40 \times 60$, and $60 \times 90 \text{ nm}^2$), materials (cobalt with $K_1 = 0 \text{ J/m}^3$, cobalt with $K_1 = 30,000 \text{ J/m}^3$, supermalloy with $K_1 = 0 \text{ J/m}^3$). The sweep was performed for each of the three latch designs (AML, SML, TML). Non-latch magnets in the system are assumed to be the same.

Note that when varying spacing, the spacing between all magnets is always the same, including between the latches. Varying the spacing of the latches only, for example, could be a subject of future work. The same is true for thickness and magnetic material. In the case of aspect ratio, only the height of the latches are changed (in the case of the AML, this means the aspect ratio of all magnets). Finally, in the case of magnet footprint, all magnets have the same width, but the latches may have increased height (depending on aspect ratio and latching method).

---

Figure 6.7. 3-phase clock for use with SML and TML. A high field is used in the first phase to place the latch in a metastable state. The field is then lowered (e.g., between 0-5 mT in this work) to clock the remaining magnets.
width and thickness as the latch, but with an aspect ratio of 1.5. Thus, for the SML and TML there is no data collected for an aspect ratio of 1.5 (as this corresponds to the AML with an aspect ratio of 1.5). These design parameters were chosen based on 1) the design space known to be reasonable due to previous work (e.g., in Ch. 5.3 [23, 24, 25, 44]), 2) the ability to make reasonable comparisons to 15 nm CMOS, 3) what can be realistically fabricated in Notre Dame’s facilities, and 4) what can serve as a good example of the methodology while requiring reasonable simulation time to investigate energy trends in Ch. 6.6. For example, spacings of 6 nm or below were not studied because it is known that 1) they will likely lead to overly strong x-directed coupling (as with the $60 \times 90 \times 30$ nm$^3$ magnets in Ch. 5.3 and as seen in [56, 55]), 2) they are difficult to fabricate, and 3) to obtain accurate results, simulations may require granularities resulting in many weeks of OOMMF simulation per design. Similarly, thicknesses greater than 20 nm were not studied because as the ratio of width to thickness decreases, the z-direction may not remain a “super hard” axis as desired (and the device may not remain single domain).

6.5.1 Methodology

In order to perform the design space sweep, the results of a large number of contour plots (e.g., Fig. 6.4) need to be captured in a succinct way. An obvious initial approach – which is pessimistic, and captures the worst possible case – is to consider all angles in the contour plots when determining the minimum stability of a configuration. (This method will be referred to as “Method 1” from this point forward.) For example, per Fig. 6.4 the minimum stability for the configuration considered (i.e., $40 \times 60 \times 20$ nm$^3$ supermalloy magnets with an 8 nm spacing) is 0 kT if one assumes that neighbor magnets in the three-magnet system can take on any angle.

However, due to thermal noise, it is unlikely that a neighbor magnet remains in a worst case state for a long period of time. Furthermore, the possible angles of the neighbors may be restricted due to the level of x-directed coupling (which depends on inter-device
spacing, etc.) between the devices. As such, if the probability of a subset of angles for the neighbors is known to be sufficiently low, those angles can be ignored in the stability comparisons [43]. (This method will be henceforth referred to as “Method 2”.) Therefore, in this work results are presented for the design space sweep that consider all possible angles (Method 1), as well as a restricted subset of angles based on OOMMF simulation results (Method 2, as proposed in [43]).

The OOMMF simulations required for Method 2 (i.e., to factor neighbor angles into the design space sweep) are carried out with lines of 15 nanomagnets, a latch positioned in the middle of the line, and an initial state that has all magnets at ±45 degrees. The magnet directly to the right of the latch begins in a conflict state and the system relaxes for 20 ns at 300 K. Upon completion, the first nanosecond of data is thrown out (to ignore the artificial starting conditions), and the minimum angle of the neighbors is observed for the remaining time (if the latch flips, it is counted as having 0 kT stability). This minimum angle dictates a cutoff for each contour, where all stability values below the minimum are ignored. The elimination of the low angles that result in high device-to-device coupling, and thus low EBs for the target, lead to higher predicted stabilities than Method 1. Due to the simulation time required for simulations at 300 K (a few days per 20 ns relaxation simulation), it is not feasible to simulate the long times needed to guarantee the neighbors will remain at these angles. Therefore, this approach may trend toward being over-optimistic, as some of the observed angles may be lower over time (due to thermal noise). As such, reality likely lies somewhere between Method 1 and Method 2, so this data is used to illustrate trends associated with design parameters of interest, which will ultimately guide future temperature-accelerated experiments to analyze the latch structures.

Due to the large amount of data, figures representing notable trends are presented, and the rest discussed in the text. First, the stability trends associated with each design parameter (spacing, thickness, aspect ratio, magnet footprint) are discussed for the TML (the AML and SML are discussed in Ch. 6.7.6).
6.5.2 Spacing

Fig. 6.8 demonstrates how stability scales as a function of spacing. In this case, latch magnet width is fixed to 40 nm, height to 80 nm, and thickness to 20 nm. (Data is presented for only the TML case, but how the AML, SML cases differ is noted.) As expected, as spacing increases, coupling between adjacent nanomagnets decreases and stability increases. Fig. 6.8 reveals an interesting trend: with Method 2, as spacing increases, there are diminishing returns with the increase in stability. For example, consider the stability of supermalloy magnets when Method 2 is employed. The stability increases 1.5X when spacing is increased from 8 to 10 nm, but this drops to a 1.10X increase in stability when spacing increases from 10 nm to 12 nm and from 12 nm to 14 nm. Finally, stability only increases by 1.02X when increasing from 14 nm to 16 nm. This suggests that increasing spacing by a small amount could be enough to provide required stability, which is highly desirable as previous work has shown that large spacings may not be viable when considering clock energy (e.g., as in Ch. 5.3.2).

6.5.3 Aspect Ratio

To show how stability scales with aspect ratio, width is fixed to 40 nm, thickness to 20 nm and spacing to 10 nm. As expected, increasing aspect ratio increases stability in all cases. The field required to null a nanomagnet increases as magnet aspect ratio increases (with spacing fixed) since as height increases and width remains fixed, $\Delta N_d$ in Eq. 6.1 increases. Results for Method 2 show, as expected, that stability is significantly higher than for Method 1.

\[\text{Note for ease of reading in grayscale: for this and all subsequent contour plots the data points listed in each legend appear from left to right in each bar graph.}\]
Figure 6.8. Stability as a function of spacing for the TML.

Figure 6.9. Stability as a function of aspect ratio for the TML.
6.5.4 Magnet Footprint

Two sizes are considered when studying magnet footprint: $40 \times 80 \times 10 \text{ nm}^3$ and $60 \times 120 \times 10 \text{ nm}^3$ (again, corresponding to sizes considered in simulation and experiment in previous work, e.g., in Ch. 5.3, [25], [2], etc.). The spacing between the devices is fixed at 10 nm, and they share the same aspect ratio and thickness. When scaling magnet footprint, the TML is again considered. Fig. 6.10 shows how stability is affected by the footprints of the nanomagnets. If Method 1 is used, the minimum stability is very low for all magnet sizes, and decreases for the larger footprint. Per the formulas in [4], the fields produced by a nanomagnet are proportional to the dimensions of that nanomagnet. Therefore, large devices will generate high fields on their neighbors. The field required to null a nanomagnet decreases as magnet size increases (with spacing fixed), since the field required to null a nanomagnet is proportional to the demagnetizing factor of the magnet, and the demagnetizing factor decreases as height/width increase (see Eq. 6.1, [4]). Thus, in the context of the 3-magnet system, larger neighbors at low angles will have a stronger nulling contribution on the target, which is in turn easier to null.

When using Method 2, increasing the size of the magnet results in a decrease in stability except for magnets with biaxial anisotropy. At high angles, $H_y$ is stronger for large nanomagnets, and they will reinforce each other’s state. Thus, for magnets with biaxial anisotropy at the larger footprint, even though the field required to null a magnet is higher, the $y$-directed coupling between adjacent nanomagnets may result in increased stability (Ch. 6.5.6 explains how biaxial anisotropy leads to higher angles due to the shape of the energy landscape). Note that the discussion thus far assumes a fixed thickness, and height/width are thus higher relative to thickness for the larger footprint devices. If thickness is scaled as well as height/width, then the relation between height/width and thickness will be the same between the two device sizes (and thus the demagnetizing factor will be the same). However, note that the stability trend is actually similar if magnet thickness is scaled along with the height/width (i.e., if one considers $40 \times 80 \times 20 \text{ nm}^3$ vs.
Figure 6.10. Stability as a function of magnet footprint for the TML.

60 × 120 × 30 nm³ magnets). This can be explained by the fact that larger magnets have more volume, and thus couple more strongly (leading to decreased stability). Therefore, in terms of stability large magnet sizes may be desirable (depending on material) if the minimum observed neighbor angles are consistent over long time periods.

6.5.5 Magnet Thickness

Next, the trend associated with scaling magnet thickness is examined. The 40 × 80 nm² magnet footprint with a spacing of 10 nm is considered. For all cases, increasing thickness results in an increase in stability. Again, this can be explained by considering the analytical formulas [4]. Per Eq. 6.1, thicker magnets have a higher demagnetizing factor and thus require a higher external field to align them along the hard axis. Furthermore, increased thickness results in increased coupling between the nanomagnets, so stability is again lower when using Method 1.
6.5.6 Material

This discussion of stability trends is concluded by considering material. First, note that the materials chosen for this study are representative of the common materials in the published literature for NML (e.g., [3, 2, 83]). Supermalloy and cobalt are materials commonly used in published research, while [13] introduced magnets with a biaxial anisotropy to preserve hard-axis stability during ensemble re-evaluation (see Ch. 5.6 for more details and related work). Per Fig. 5.4, adding biaxial anisotropy flattens the peak of the energy landscape for a nanomagnet and can introduce a local energy minimum at 0° (depending on the value of the biaxial anisotropy term, K1). For example, in this study cobalt nanomagnets with K1 = 30,000 J/m$^3$, and K1 = 0 J/m$^3$ are considered. In Fig. 5.4 the curve for K1 = 30,000 J/m$^3$ has a local energy minimum at 0°, while the curve for K1 = 0 J/m$^3$ has none. In [13], the local energy minimum is introduced in the context of soliton switching (discussed in more detail in Ch. 5.6.2) to ensure that magnets remain hard-axis biased until
driven to a logically correct state by their neighbors. In the context of this work, magnets with biaxial anisotropy could have beneficial properties. For example, they could lead to lower power consumption, as described in Ch. 5.6, but may also lead to higher stability.

In the context of stability trends, when Method 2 is used, ensembles of supermalloy devices are the least stable in all cases, followed by cobalt, and then cobalt with $K_1 = 30,000 \text{ J/m}^3$ (e.g., see Fig. 6.8). However, when Method 1 is used, cobalt with $K_1 = 30,000 \text{ J/m}^3$ is the least stable, followed by supermalloy, and finally cobalt with $K_1 = 0 \text{ J/m}^3$. This result suggests observing the neighbor angles for longer simulation duration (i.e., in Method 2) may be most important for cobalt with biaxial anisotropy. That is, if the minimum observed angle for cobalt with $K_1 = 30,000 \text{ J/m}^3$ is significantly lower over long simulation time there could be a large impact on calculated stability. The difference in trends between Methods 1 and 2 can be explained by comparing the energy barrier curves for biaxial and non-biaxial cobalt nanomagnets. As seen in Fig. 5.4, the slope of the energy barrier is steeper for magnets with biaxial anisotropy. Thus, at higher angles of magnetization (i.e., closer to $\pm 90^\circ$), the magnets with biaxial anisotropy will have higher stability (i.e., the energy difference between lower angles of magnitude, e.g., between $+90^\circ$ and $+45^\circ$, etc. will be larger). This leads to higher stability when Method 2 is considered as many of the lower angles of magnetization in the contour plots are not counted in the stability calculation.

6.6 Energy Trends

In this section, results are presented from the study of energy requirements for the design space defined in Ch. 6.5. In order to examine energy trends associated with the design parameters of interest, OOMMF simulations are run to determine the minimum required field strength to clock each design. Simulation methodology is described first followed by a discussion of energy requirement trends associated with each design parameter.
6.6.1 Methodology

Simulations to study energy requirements were again carried out with OOMMF using lines of 16 nanomagnets terminated with a block of magnetic material. The line is broken into three groups of 5 nanomagnets to operate with the 3-phase clock described in Ch. 5.1.1. The first magnet is unclocked and fixed at 45° to act as an input to the line of nanomagnets. The terminating block is also unclocked and biased in the positive x-direction to manage the boundary condition (as in Ch. 3.1.2, Ch. 5.2.3, etc.). In the case of the TML and SML, the latches are placed at the start of each clock group and the waveform described in Ch. 6.3.2 and Fig. 6.7 is used. Propagation of the input signal is considered to be successful if in-order switching from input into the last clock group is observed. Multiple simulations for each design were performed to determine the minimum required field (with granularity of 5 mT) to facilitate complete signal propagation (tϕ was fixed at 3ns for this study, but this could be varied in future work). All simulations are performed at 300 K using OOMMF.

For the sake of comparison, energy is estimated by the square of the clock field (in mT) in each phase required for propagation (since the required current scales proportionally with applied field, and the phase time is fixed). Therefore, in each plot energy is presented as $H^2$ to show relative energy between configurations (this also takes into account the energy from the unique waveforms for SML/TML as described in Ch. 6.3.2 and Fig. 6.7). In reality, the energy will change slightly based on the resistance of the clock wires. Since power and energy are dominated by $I^2$, the slight differences in resistance are ignored when comparing energy. If one were to calculate energy directly, it would take a significant amount of overhead per design. First, Maxwell clock wire design must be created to match the width of a clock wire group for each design (this requires drawing the geometry of the wire by hand in software). Next, multiple simulations are run at different currents to generate the desired field strengths. Once the field distributions are generated, data is extracted and imported into the OOMMF simulation setup. This process is clearly
prohibitive, given the large number of simulations in this design space study.

In the interest of time, only the designs associated with Figs. 6.8-6.11 are simulated.

6.6.2 Spacing

First, note that at an 8 nm spacing the cobalt magnets with biaxial anisotropy could not be clocked. At approximately 35 mT, the clock field is sufficient to null each group of magnets, but magnets in the second clock group relax prematurely, leading to logically incorrect propagation. This result is not surprising, as previous work has shown that these devices typically work best with higher inter-device spacings.

As expected, clock field requirements increase with spacing. As seen in Fig. 6.12, increasing spacing results in a large increase in field requirement past 10 nm (higher than 100 mT). This is likely a function of the magnet footprint – that is, 60 × 120 nm$^2$ devices will likely operate at a relatively low (i.e., below 100 mT) field at 12 nm spacing (suggested by results in Ch. 5.3.2). That said, above 12 nm the larger magnets will likely follow the same trend when spacing is increased further. Thus, larger spacings could be a subject of future study but is beyond the scope of the design space considered here. Note that initial simulations considered spacings larger than 12 nm, but results readily demonstrated that the required clock field increased rapidly at these spacings (also, the stability was subject to diminishing returns as spacing increased, as described in Ch. 6.5.2). In order to reduce simulation time, these designs are not pursued further, and stability data is not presented.

6.6.3 Aspect Ratio

Fig. 6.13 shows how energy scales as a function of aspect ratio. As aspect ratio increases, energy quickly goes up as well. Depending on the material, each increased step in

---

4 The simulation time for a single design is approximately 15 hours (each simulation uses 4 cores), and approximately 10 simulations are required for each design to determine the minimum field strength. Thus, to simulate the entire design space it would take approximately 11 years of simulation (or approximately 200 days if 20 jobs are constantly executing).
Figure 6.12. Energy (estimated by mT^2) as a function of spacing for the TML. Note that at an 8 nm spacing the lines of cobalt magnets with K1 = 30,000 J/m^3 could not be clocked.
AR can result in up to twice the energy requirement. It is likely difficult to achieve fine-grained fabrication control of aspect ratio. That is, an AR of 1.75 versus 2.0 corresponds to a 10 nm difference in height for a width of 40 nm, and achieving 5 nm resolution may be difficult. Therefore, if an AR of, for example, 1.8 is desired, even a small variation could result in a fairly large impact on field requirement (or result in defects).

6.6.4 Magnet Footprint

Energy results appear in Fig. 6.14. Note that the larger magnets (60 × 120 nm²) require the least amount of clock field at this spacing and thickness. Therefore, in terms of energy, larger magnets may be more desirable if area is not a concern. For applications where scaling is desired, the 40 × 80 nm² footprint is viable, but might come at the cost of increased energy.

Note that while magnets with a footprint of 20 × 40 nm² were studied, these lines were too susceptible to thermal fluctuation and the signal would not propagate (stability data was thus not presented). However, this does not indicate that this magnet footprint is unusable, but that it is not viable for this specific design space. In order to make an “apples to apples” comparison to the larger footprints, both the spacing and thickness of the devices was fixed to 10 nm. It is possible that lines of 20 × 40 × 10 nm³ devices could operate correctly at closer spacings, or with a different thickness or aspect ratio.

6.6.5 Magnet Thickness

Results are presented in Fig. 6.15. In contrast to the energy trends associated with aspect ratio and spacing, increasing magnet thickness results in only a minor energy cost. Note that devices with a thickness of 5 nm were initially considered. However, thermal fluctuation restricted signal propagation and in-order operation. Thus, results are not presented for energy (and were not presented for stability). However, this does not rule out the use of thin devices – for different spacings, magnet footprints, aspect ratios, etc., thin
Figure 6.13. Energy (estimated by mT$^2$) as a function of AR for the TML.

Figure 6.14. Energy (estimated by mT$^2$) as a function of magnet footprint for the TML.
6.6.6 Material

Energy trends for the three materials of interest can be understood by considering Figs. 6.12-6.15. In all cases, the material that requires the highest energy is cobalt with $K_1 = 0 \text{ J/m}^3$. For the other two materials (supermalloy and cobalt with $K_1 = 30,000 \text{ J/m}^3$), there is not a clear “winner” in terms of energy. In most cases, the two materials require roughly the same energy. However, at the extreme ranges of the design space for spacing (12 nm) and aspect ratio (2.25), supermalloy requires significantly higher energy.

6.7 Stability vs. Energy Trends

This section discusses trends to help show which designs can provide high stability and low energy. The metric $kT/mT^2$ is used to show the amount of stability provided per required energy. High values are obviously desired as they correspond to designs that offer high stability with low energy requirements. As an example, a 5 mT clock field has been shown to allow NML to compete with low-power CMOS (in Ch. 5.3), and 40 kT is a commonly desired stability. Thus, achieving a $kT/mT^2$ of 0.8 ($40/(5^2 + 5^2)$ to account for the two phases where field is applied) would be a near-ideal outcome, and will be a measuring point to refer to going forward. Recall that energy is estimated by $H^2$ since the current in the clock lines is proportional to the field generated, and energy is proportional to the square of the current. Again, results are presented using the TML, and discussion is concluded with trends for other latching methods.

6.7.1 Spacing

Fig. 6.16 depicts stability/energy as spacing is scaled. A spacing of 8 nm offers the lowest stability, so stability/energy is low despite the modest energy requirements. Similarly,
given a 12 nm device-to-device spacing the high stability is offset by the high energy requirements. Therefore, due to the high stability/energy, a 10 nm spacing is most desirable. Furthermore, the stability for a 10 nm spacing is significantly higher than the typically-desired 40 kT even in the worst case (for all materials). The takeaway from this study is that there is likely a “sweet spot” for spacing – specifically, note how kT/mT^2 is highest when spacing is 10 nm (though never more than half of the desired 0.8 per Ch. 6.7). Furthermore, this suggests fabrication tolerances may be important to ensure that spacing is well-controlled (otherwise, large swings in stability and energy could occur).

6.7.2 Aspect Ratio

Fig. 6.17 depicts stability/energy as aspect ratio is scaled. It is clear that the best choice for aspect ratio is 1.75. Results indicate that an aspect ratio of 1.75 offers high stability while also requiring low energy to be clocked. As a result, stability/energy is significantly
Figure 6.16. Stability (kT) per energy (estimated by mT^2) as a function of spacing for the TML. Recall that for an 8 nm spacing, the lines of cobalt magnets with K1 = 30,000 J/m^3 could not be clocked.
higher. Specifically, $kT/mT^2$ is 0.6 when considering Method 2 (this begins to approach the desired 0.8 per Ch. 6.7). However, in the worst case, stability/energy is quite low for an aspect ratio of 1.75 (due to low stability). Future experimental efforts will determine if this aspect ratio can provide the high stability suggested by the relaxed model.

6.7.3 Magnet Footprint

The stability/energy trend for the two magnet sizes studied presents an interesting design tradeoff. The results are depicted in Fig. 6.18 and clearly show increased stability/energy for the larger magnet footprint. (However, note how $kT/mT^2$ is significantly lower than the desired 0.8 per Ch. 6.7 due to the high aspect ratio chosen to illustrate this trend.) This introduces a tradeoff between area, stability and energy. The choice of small or large magnet footprint will likely depend on the application of interest and material considerations (e.g., whether magnets with high biaxial anisotropy can be fabricated). For example, an application where non-volatility and low energy are stressed could sacrifice device footprint scaling.

6.7.4 Magnet Thickness

Increasing thickness seems to be the most promising avenue to creating high stability, low energy devices. Fig. 6.19 shows how stability/energy goes up quickly as thickness is increased. It is intuitive that thicker magnets will be more stable, but somewhat counter-intuitive that it can lead to decreased energy. For a magnet in isolation, increased thickness should result in higher required nulling fields, but the increase in coupling between thicker devices in ensembles can result in lower energy operation (this has been observed in previous work [44, 43]). Therefore, the most desirable thickness in this study is 20 nm. Again, $kT/mT^2$ is highest (over 0.15) when the 20 nm thickness is chosen – note again that though this is much lower than the desired value of 0.8 per Ch. 6.7 this is due to the choice of aspect ratio to illustrate this trend. Future work should look at the maximum thicknesses
Figure 6.17. Stability (kT) per energy (estimated by mT$^2$) as a function of aspect ratio for the TML.
possible for designs of interest. However, there is a clear upper limit as it is undesirable for the z-direction to become an easy axis of magnetization.

6.7.5 Material

Each of Figs. 6.16-6.19 shows the stability/energy trends for different materials. For Method 2, it is clear that cobalt with $K1 = 30,000$ K J/m$^3$ provides the greatest stability/energy. In fact, the highest $kT/mT^2$ is achieved when this material is chosen (specifically, 0.6 in Fig. 6.17). This trend is supported by simulation, and is actually encouraging as it seems these devices could be stable when the clock is removed, but respond to clock fields in a way that results in lower energy requirements.
6.7.6 Latching Method

The discussion of stability and energy trends is concluded with a comparison of the three latching methods. Results show that when Method 1 is used, the TML is the most stable, followed by the SML, and finally the AML (for all materials). When Method 2 is used, the AML is the most stable latching method, followed closely by the TML and finally the SML. This is intuitive, as in the AML approach the energy barrier of all magnets is raised (due to higher aspect ratio, or decreased coupling due to spacing, etc.) which leads to neighbor angles near ±90°. In the other two cases, the unstable 40 × 60 × 20 nm³ magnets in the line can relax to low angles and in turn lower the angles of the latch, which can transition through an F-ordered state (as in Fig. 6.2). This effect is strongest on the SML because it is directly adjacent to two 40 × 60 × 20 nm³ neighbor magnets, whereas the center magnet in the TML is only indirectly impacted.

In terms of energy requirements, the TML again seems to be the most promising.
Though the stability provided can be higher than the AML, the required energy to clock the TML is significantly lower. The reason for this is that the TML can take advantage of the unique waveform in Fig. 6.7 in which the clock field in the “switching phase” is reduced to between 0-5 mT (i.e., the field required to drive the signal through the $40 \times 60 \times 20$ nm$^3$ non-latch devices). The SML similarly benefits from this unique waveform.

Furthermore, when considering each design parameter (e.g., spacing), the stability and energy trends are similar for all three latching methods. Therefore, since the TML is likely the most stable, is lower energy, trends are similar, and in the interest of space, trends with the other two latching methods are not discussed in detail.

### 6.7.7 Summary and Future Work

Recall that due to the limitations of the analytical method used in this work, one cannot predict the absolute stability of any of these designs. Rather, designs are demonstrated that offer improved stability (with the least energy cost) over the un-latched case (which represents a design central to NML research efforts). Ultimately, the absolute stability of the designs will need to be determined to, for example, determine which aspect ratio, spacing, etc. is sufficient to offer a desired stability (dictated by the application of interest). Theoretically, this could be accomplished using micromagnetic simulations, but the run-time of simulations to gather meaningful data is prohibitive. In order to gather this data, temperature-accelerated experiment can be leveraged, e.g., by observing bit flips of a circuit in a 700 K environment over short timescales, one could extrapolate to stability in a 300 K environment for a longer-duration (e.g., as was done for MRAM in [1]). This work demonstrates how various design parameters affect stability and energy, and can thus be used as a guide for these future experimental efforts.

In general, this work has shown that experimentalists should first study designs with relatively low-AR, low-spacing, high-thickness magnets with biaxial anisotropy and a three-magnet latch. Though none of the designs reach a kT/mT$^2$ of 0.8 as described in Ch. 6.7.
(which would be ideal, or nearly so), this work has identified designs that should be studied first in experiment. For example, the most obvious design to study in experiment is a TML of $40 \times 70 \times 20$ nm$^3$ devices made of cobalt with $K_1 = 30,000$ KJ/m$^3$ spaced 10 nm apart. This design achieves a $kT/mT^2$ of 0.6, the highest of any design considered. However, a high $kT/mT^2$ could be achieved with an unrealistic field strength. In this example, the required field is 40 mT. To see how this could impact system-level energy, recall the comparison of an NML RCA to a low-power CMOS equivalent in Ch. 5.4.1. As an estimate, increasing clock field strength from 5 mT to 40 mT would require 1,575X higher energy. Recall the best NML performance (the last row in Table 5.4) resulted in 128X improvements in terms of EDP. Thus, this design would yield approximately 12X worse EDP than CMOS equivalents. However, it is important to stress that this is but a starting point, and the tradeoff between non-volatility and low energy will depend heavily on the application. In other words, the CMOS equivalents are volatile, and thus a 12X worse EDP might be acceptable depending on the application (again, further study is needed to determine ideal designs).

Though it seems the cobalt magnets with biaxial anisotropy could provide the highest $kT/mT^2$, supermalloy has been considered in great depth, and proven methods exist to fabricate NML circuits with this material. As such, though supermalloy is not necessarily the best choice to offer high stability and low energy, it should also be targeted in experiment to facilitate characterization of ideal designs. Finally, if area is not a priority, one could increase the footprint of the magnets, e.g., to consider $60 \times 105 \times 20$ nm$^3$ devices. Larger footprints may be desirable as they could require less clock energy and provide higher stability (depending on material). Both device footprints should be considered in experiment, and a thorough analysis of the tradeoffs of scaling on area, throughput, etc. should be made for applications of interest (e.g., in the context of the checkpointing application to be discussed in Ch. 6.8).

These two designs are by no means all that should be studied, but with the expense of
experiment (both in time and materials) they represent a good starting point suggested by this work. For future NML design efforts that might be focused on a different design space, the format of this study can be followed. That is, the analytical method from [43] can be used to evaluate trends in stability and energy, which can be used to identify promising designs to study in experiment. The tradeoffs between stability and energy must ultimately be weighed in the context of the application of interest and the experimental results.

6.8 Potential Shift Register Application

Now that it has been shown how to provide non-volatility for NML shift registers, this section discusses potential applications for a non-volatile NML shift register. First, an example design is presented that is used to make a comparison to two proposed checkpointing applications.

6.8.1 Checkpointing

A potential application for the NML shift register is as a memory for checkpointing. If the shift registers are non-volatile, they can store computational state that can be rolled back in the event of a power failure or soft error. A main goal of a checkpointing memory is to reduce the amount of processing time that is lost due to power failure or soft error. Checkpointing also enables instant-on operation as state can be restored from the memory rather than disk. When designing a memory used for checkpointing, one must consider the following in the context of minimizing lost processing time:

- the checkpointing and recovery operations must be as fast as possible;
- the time between checkpoints must be sufficiently frequent.

Methods exist for optimizing the checkpoint interval given the time required to checkpoint, the recovery time, etc. For example, a first order approximation of the optimal time between checkpoints is given in [88], and a higher-order approximation in [20]. These and
other methods should be considered in order to design the shift register to optimize performance. That is, the first step will be to optimize a single shift-register cell to determine the trends associated with checkpoint and recovery latency, optimal checkpoint interval, etc. However, as will be discussed in Ch. 6.9, these considerations will ultimately depend on I/O energy and latencies, which are not well-characterized, and thus will be left to future work.

6.8.2 An Example Design

First, an example shift register memory cell design is discussed to use as a basis for comparison to other approaches. Note that this example is not meant to describe an ideal design but rather to demonstrate a design that could be realistically implemented and can compete with existing technologies.

The bit density is simply a function of the magnet dimensions and the spacing between magnets. The area per bit is therefore defined as:

\[ n(x + s_x) \times (y + s_y) \]  

(6.2)

where \( n \) is the number of magnets per bit, \( s_x \) and \( s_y \) are the spacing between magnets in the x- and y-directions, respectively, and \( x \) and \( y \) are the magnet width and height, respectively. \( n \) is defined by the number of magnets per clock group, and the number of clock groups per bit (\( \phi \) groups per bit are required in a \( \phi \)-phase clock). With a 3-phase clock scheme it will take three clock groups to represent a single bit. Assuming a 3 ns phase time as in Ch. 5.3 and Ch. 6.6 the time to shift one bit from input to output of a single shift register will be 15 ns. Another way to think about this “shift time” is the delay between subsequent writes to the same shift register.

Recall that shift registers can be designed to be volatile, or non-volatile. A volatile shift register is a necessary component in a variety of circuits, e.g., the systolic convolution
application described in [18]. Consider a shift register with parameters similar to those in Ch. 5.3: supermalloy magnets; magnet size of $40 \times 60 \times 20 \text{ nm}^3$; inter-device x-spacing of 8 nm; 60 nm y-spacing between adjacent shift registers; a 3-phase clock for the shift operation; 10 magnets per clock group (i.e., per clock wire). Performance of lines with these parameters has been well-studied. The area per bit for this example would be 0.1728 $\mu\text{m}^2$ per Eq. 6.2. Assuming the 3-phase clock and 3 ns phase time, the energy to shift in a single bit to the shift register is 424 aJ. Energy is calculated with the equations from Ch. 5.1 as in Ch. 5.4 and thus includes energy dissipated in the clock wires, clock driver overhead, etc.

A non-volatile shift register could also be useful in a variety of applications, e.g., in a checkpointing application as will be discussed in Ch. 6.8.1. As absolute stability cannot be predicted (per Ch. 6.4.1), an “increased stability” shift register is considered instead. Future experimental work is needed to determine whether this design will provide sufficient stability (and it will depend on the desired stability time for the application). That said, as an example, consider a shift register with the following parameters: a three-magnet latch; supermalloy magnets; magnet size of $40 \times 70 \times 20 \text{ nm}^3$; inter-device x-spacing of 10 nm; 70 nm y-spacing between adjacent shift registers; a 3-phase clock for the shift operation; 10 magnets per clock group (i.e., per clock wire). The increased stability of this design is discussed in Ch. 6.5 while the required clock field increases from 5 mT to 35 mT. Recall that the TML can take advantage of the waveform in Fig. 6.7, so energy increases 25X (rather than 49X). The area for this design is 0.21 $\mu\text{m}^2$ per Eq. 6.2 and the energy per shift (following the methods in Ch. 5.1) is 11,200 aJ (approximately 25X higher than the volatile shift register).

The read latency and energy requirements for a single shift register will be those of the output structures discussed in Ch. 2. Similarly, the shift times and shift energies need to be augmented by input latency and energy. The energy and delay requirements for the NML I/O approach presented in Ch. 2 have not yet been well-characterized, but they
are expected to be significantly lower than similar metrics for the MRAM equivalents. Performance savings are possible as the I/O devices can be hard-axis biased when set. For example, if the input device is hard-axis biased the current required for input should be small (the device just needs a small bias to be tipped). The read/write times for MRAM are on the order of tens of nanoseconds, with a write energy of 50 pJ per bit [71]. Alternatively, if a bias line is employed, per [80], the energy per input event is approximately 0.7 fJ if a 1 ns write time is assumed (which per [80] is a reasonable assumption). For the sake of this discussion, these latencies and energies are assumed, but the MRAM latencies and energies are also considered as a worst case, upper bound approximation.

Finally, note that shift registers can be chained horizontally so a single line represents multiple bits, e.g., a line of N groups can represent N/\(\phi\) bits (where \(\phi\) is the number of phases in a \(\phi\) phase clock). Long lines can therefore share the same I/O, reducing the number of magnetic-electrical interfaces (MEIs) that are needed. There is thus a tradeoff between the number of MEIs and the granularity of access as a line of many bits could see significant delay when shifting from input to outputs. This design parameter will ultimately be tuned based on the application’s requirements.

6.8.3 Comparison

Now, comparisons are made to existing approaches. NML needs to offer better performance in one or more ways (e.g., energy, delay, area, etc.) in order to be useful in a checkpointing architecture. Due to the nature of NML devices and the study presented in Ch. 5.4.1 the goal for NML checkpointing is namely to realize benefits in terms of low power operation, non-volatility, and decreased delay (when compared to writing to non-volatile disk-based storage).

The checkpointing methods proposed in [28] and [42] have similar goals. [28] proposes a hybrid local/global checkpointing for massively parallel processing (MPP) systems that uses phase change random access memory (PCRAM) devices to store checkpointed data.
These systems currently use disk-based checkpointing methods, which become prohibitive as the number of processing nodes increases to the exascale and beyond. The goal in [28] is therefore to offer improved performance compared to disk-based approaches. The PCRAM device consists of a phase-change material that exhibits high resistivity when in an amorphous state, and low resistivity when in a crystalline state – these two material states can be used to represent binary state. State transition, i.e., a write operation, is achieved by heating the material by applied electric current. The two states can be read by sensing the resistance of the material. PCRAM is non-volatile, has no standby power, and should have relatively low read latencies.

[28] proposes to develop PCRAM chips that can be integrated into a dual in-line memory module (DIMM) for use in MPP systems. An NML-based memory for use in this application thus would be fabricated in chips that could be arranged in DIMMs. If an NML shift register cell can provide lower energy, read/write latencies, and density then it could replace PCRAM memory cells for this application. [28] reports the phase change device to have read times of 10-100 ns, write times of 100-1000 ns, and a write energy of 50-100 pJ. The 0.7 fJ projection for input energy and 1 ns projection for latency in Ch. 6.8.2 compare well to those in [28]. Furthermore, even the worst case approximations associated with MRAM I/O (50 pJ write energy per bit, tens of nanoseconds write latency) outperform the PCRAM estimates. Also, recall these worst case approximations can be relaxed if one can leverage performance improvements as suggested in Ch. 6.8.2 (e.g., hard-axis biasing input magnets). In either case, the delay and energy per shift operation for the improved-stability shift register are 15 ns and 11,200 aJ (i.e., 0.0112 pJ), respectively, so are dwarfed by the I/O performance. These initial comparisons suggest that NML could perform similarly to or better than PCRAM devices for this application, and thus merit further exploration.

[42] proposes a checkpointing architecture that integrates non-volatile devices to checkpoint at the micro-architecture level. The goals in this approach are to protect the processor from lost computation time, as well as to allow for instant-on operation. The hybrid Hall
The Hall effect (HHE) device consists of a small ferromagnet (thus, states are encoded in magnetic polarization as in NML), a wire that runs below the device, and a wire that runs above the device. The two wires are orthogonal; the top wire allows for writing (via current-induced fields from the wire), while the bottom wire allows for reading via the Hall effect. The Hall effect states that when a current is applied to the lower wire, the magnetic field of the ferromagnet will create a voltage in the wire orthogonal to the direction of current flow. The sign of this voltage depends on the orientation of the ferromagnet, and thus can be sensed to determine device state.

The proposed architecture in [42] requires that all micro-architectural state – e.g., the register file, program counter, etc. – be preserved by the non-volatile checkpointing devices. Therefore, each bit of these structures would need to interface with an NML shift-register cell. The method proposed in [42] is to have an HHE device coupled with each bit. Thus, during each checkpoint or restore operation all saved state could be written or read in parallel. The natural approach would be to have one NML shift register per bit. Therefore, if NML can provide lower energy, delay and area it can be a viable replacement in the same architecture. However, there are other architectural considerations unique to NML, as multiple bits will likely share the same clock wire. Therefore, the optimal configuration of the NML shift registers in terms of bits per clock wire, and any other similar considerations that arise, should be studied. Again, this is left to future work as it will ultimately depend on I/O characteristics (see Ch. 6.9). For the time being, comparisons will be made on the level of a single memory cell in order to show that NML may offer improved performance.

The read and write times of the HHE device are 500 ps as reported in [42]5 with possible improvements allowing for reductions down to 100 ps. These projections are significantly faster than both the worst case assumption (MRAM read/write times of tens of nanoseconds) and the estimated bias-line write time (1 ns). However, if increased performance for NML can be achieved by biasing I/O devices along their hard axes, perfor-

---

5Note that the authors state that read times have not been well-studied, but project 500 ps.
mance improvements could potentially be realized. Since the checkpointing interval is much larger than device write times (thousands of cycles in [42]) the time to shift from inputs to outputs – again, 15 ns based on the example in Ch. 6.8.2 – is of concern only if power is lost during the shift operation. [42] does not explicitly state read and write energies for the HHE devices, but [39] reports power requirements of 0.1 mW for devices similar in dimensions to the baseline in [42]. This power consumption is significantly higher than that of NML clock wires – on the order of µW per clock wire excitation (see 4.2.1) – which are the dominant source of power consumption in NML circuits. Though more work is needed to compare the expected energies of the two technologies, this data point suggests that NML could be quite an attractive device to replace the HHE device in this architecture.

Finally, as the HHE approach requires that all architectural state be preserved any device used in this architecture must not add significantly to the area requirements, as it will be fabricated on-chip. The features of devices studied in [42] are on the order of a few microns, which is significantly larger than the dimensions of one shift-register cell if made of $40 \times 70 \times 20 \, \text{nm}^3$ devices (.21 $\mu$m$^2$ as projected in Ch. 6.8.2). [42] projects the HHE devices to scale with the feature size of manufacturing processes, and scaling limitations associated with the ferromagnets will apply to NML as well. Future study is warranted as the possibility for improvements in terms of energy, delay, and area make NML shift registers a useful replacement for HHE devices in this checkpointing scheme.

6.9 Conclusions and Future Work

I have demonstrated that NML shift registers work in simulation and can be designed to be non-volatile at the cost of increased energy requirements. Future experimental work is needed to validate designs, but I have shown that a NV-design could be significantly more stable than the initial, volatile design. I have also reviewed a methodology to quantify stability trends [43], described a methodology to quantify energy trends, and used these
methodology to show trends for a design space of interest. Notably, I identified promising designs that could offer high stability with a relatively low increase in required energy (which should be targeted first in experiment). Specifically, the most obvious design to study in experiment is a three-magnet latch consisting of $40 \times 70 \times 20$ nm$^3$ devices made of cobalt with biaxial anisotropy ($K1 = 30,000$K J/m$^3$) spaced 10 nm apart. This design achieves a $kT/mT^2$ of 0.6, the highest of any design considered, and is a starting point to guide temperature-accelerated experiment to work toward identifying ideal designs.

I have also shown that NML shift registers could compete with PCRAM in the general checkpointing application proposed in [28], or provide improved performance in the low-level checkpointing application proposed for the HHE devices in [42]. However, the performance of NML I/O will ultimately determine whether performance wins can be realized. If the energy, delay, and area of NML I/O are each significantly less than for the shift operations, then it should be possible to achieve improvements over the two checkpointing applications. Even if this is not the case, one could for example, amortize area penalties from I/O by increasing the number of bits stored horizontally in each shift register. However, without a full characterization of NML I/O at the time of this writing, it is unknown in which direction to go in terms of layout, etc. Therefore, though NML shift registers in a checkpointing application seem to be promising, a full characterization is left to future work when I/O is developed.
CHAPTER 7

STOCHASTIC COMPUTING WITH NML

The work throughout this document represents an effort to elevate the understanding of NML from the level of devices to circuits, and then from circuits to applications and architectures. This allowed me to make projections for NML to compare to CMOS in Ch. 5. Furthermore, in Ch. 6 I showed how to identify designs for non-volatile shift registers with low energy requirements, as well as how NML shift registers could compete with CMOS equivalents for an application of interest (checkpointing). Previous work has shown how an NML device architecture maps well to certain system architectures, e.g., a systolic architecture in [41] and [18], or a non-Boolean architecture as shown in [63]. NML maps well into these architectures as they take advantage of the positive aspects of NML (e.g., non-volatility, low energy, inherently pipelined data flow), while mitigating the negative aspects (e.g., high delay as compared to CMOS). Therefore, in this chapter, I present efforts to determine how NML might be used in a stochastic computing architecture, which might similarly take advantage of NML’s benefits, while mitigating drawbacks.

First, in Ch. 7.1 I describe the principles of stochastic computing, including methods to implement stochastic logic (Ch. 7.1.1-7.1.2). Next, in Ch. 7.2 I describe the principles of perpendicular magnetic logic (PML) that will be employed to make comparisons to CMOS equivalents. In Ch. 7.3 I describe two circuits used to compare performance between a stochastic PML implementation and deterministic and stochastic CMOS implementations. I conclude in Ch. 7.4.
7.1 Stochastic Computing

Stochastic computing performs Boolean logic operations on random bit streams of data \[33, 32\]. This is in direct contrast with the majority of current logic that operates on deterministic data. In stochastic logic, a value \(x\) between 0 and 1 is represented by a stream of random bits. For example, consider a number \(x\) (between 0 and 1) represented by 10 bits. The probability of each of the 10 bits being a 1 is \(x\). Therefore, to represent \(x = \frac{7}{10}\), 7 of the bits in the stream of 10 bits are expected to be 1, and the rest 0. Thus, 1111111000 or 1010101111, etc. could represent \(\frac{7}{10}\) as each stream of bits has \(\frac{7}{10}\) of its digits set to 1, and the rest 0. When these bit streams are fed to stochastic logic elements, they can be arranged in serial, or in parallel. When arranged in serial, the bits shift through the stochastic processing elements one after another. When arranged in parallel, multiple copies of the stochastic processing elements operate on each bit in the bit stream.

Stochastic logic is best illustrated with an example. Consider Fig. 7.1a-b which depicts how a stochastic multiplier would operate on the bit stream of length 10. Note that with 10 bits, one can only represent numbers \(\frac{1}{10}\), \(\frac{2}{10}\), etc. (thus, stochastic logic is less compact than conventional logic as it takes a stream of \(2^M\) bits to achieve a precision of \(\frac{1}{2^M}\)). Fig. 7.1a-b shows an example of multiplying \(\frac{5}{10}\) and \(\frac{2}{10}\). As expected, the result is \(\frac{1}{10}\). (Thus, if the probability of obtaining a 1 at the inputs is \(p_{in1}\) and \(p_{in2}\), then the probability at the output is \(p_{out} = p_{in1}p_{in2}\).) Note the difference between the serial and parallel implementations: the former sees bits shifted in over time, while the latter computes the result in a single timestep with multiple copies of the logic. Finally, consider Fig. 7.1a-b in the context of bit-level errors and fault tolerance. One of the major benefits of stochastic logic over deterministic logic circuits is its ability to withstand errors. For example, if one bit in each of the input streams in Fig. 7.1a-b is randomly flipped, the worst case result is \(\frac{0}{10}\) or \(\frac{3}{10}\) at the output. Each bit flip at worst changes the output by \(\frac{1}{10}\). However, in the case of a conventional multiplier, a bit flip at the most significant bit of either of the inputs would result in significantly higher error (e.g., up to \(2^M\) if the most significant bit, M, is flipped).
Figure 7.1. (a) Serial implementation of stochastic multiplier via a single 2-input AND gate. (b) Parallel implementation of stochastic multiplier via 10 2-input AND gates. (c) Serial implementation of stochastic scaled addition via multiplexer.
Next, another of the basic building blocks in stochastic logic is considered: the multiplexer. In stochastic logic, the multiplexer implements *scaled* addition of inputs A and B by a select signal S, i.e., \((SA) + (1-S)B\). As an example, consider the (serial) multiplexer in Fig. 7.1c: When S is 0, input A is selected as the output, and when S is 1, input B is selected as the output. With the inputs A, B, and S as \(\frac{4}{10}\), \(\frac{2}{10}\), and \(\frac{5}{10}\) respectively, the output is \((\frac{5}{10} \times \frac{4}{10}) + (1-\frac{5}{10}) \times \left(\frac{2}{10}\right) = \frac{3}{10}\). Other simple logic gates can also be used: an OR gate implements the function \(A + B - AB\) (for inputs A, B), while an inverter implements \(1 - A\) (for input A).

Stochastic computing offers other benefits in addition to a high tolerance for errors. First, stochastic computing can be more compact than deterministic implementations due to the simplicity of certain logic operations (e.g., a stochastic multiplier is a single AND gate, and scaled addition requires a single multiplexer). Due to the nature of computing on random streams of bits, stochastic computing lends itself to implementations of certain logic functions. A common example is for image processing functions, which can be implemented directly [52], or indirectly via approximation [75]. Image processing functions typically operate on large amounts of streaming data (i.e., many pixels), and are composed of simple logic functions. For example, [75] shows how to approximate the gamma correction function, and [52] shows how to implement edge detection. In both cases, error rates and hardware use are significantly lower for the stochastic implementations as compared to deterministic implementations. Before discussing comparisons to NML, the methods to synthesize stochastic logic are discussed.

7.1.1 A Method to Synthesize Stochastic Logic

Our goal is to show how NML could potentially be used to realize stochastic computing architectures. Therefore, two methods will be considered that compare stochastic NML to deterministic CMOS and stochastic CMOS logic. The first method in Ch. 7.3.1 considers comparisons for a simple multiplier, while Ch. 7.3.2 considers a more complex
comparison based on a method to synthesize stochastic logic functions (proposed in [75]). Details and mathematical proofs of this method can be found in [76], here the idea is presented at the functional level, and how their method compares to a conventional implementation is noted. With this understanding, a comparison can be made to an NML stochastic implementation in Ch. 7.3.2.

[75] has shown how to implement many polynomials with stochastic logic by converting each polynomial into a specific form of polynomial called a *Bernstein polynomial* (and ensuring the coefficients of the polynomial are on the unit interval). These Bernstein polynomials are then computed using the generic circuit described in [75]. More specifically, the method in [75] centers around the observation that stochastic logic typically implements a certain type of polynomial. These polynomials are represented in *power form*:

\[
g(t) = \sum_{i=0}^{n} a_i t^i
\]  

where all \( a_i \geq 0 \) and \( \sum_{i=0}^{n} a_i \leq 1 \). [75] shows how to convert from power form polynomials into the Bernstein polynomial form:

\[
B_n(t) = \sum_{i=0}^{n} b_{i,n} B_{i,n}(t)
\]  

where all \( b_{i,n} \) are on the unit interval, and the term \( B_{k,n}(t) = \binom{n}{k} t^k (1-t)^{n-k} \). Then, they describe a stochastic logic circuit that they show can implement any Bernstein polynomial.

An example of this circuit for a degree-3 Bernstein polynomial is shown in Fig. 7.2. The circuit consists of two blocks: a decoder block, and a multiplexing block. The decoding block consists of an adder tree and a decoder: its inputs are \( n \) independent stochastic bit streams (for \( n \) in a degree-\( n \) Bernstein polynomial). The purpose of the decoding block is to count the occurrences of 1s in each of the inputs and feed this result to the multiplexing block. For a degree-3 Bernstein polynomial, [75] shows that a full adder and 2-to-4 decoder
are required for the decoding block. The multiplexing block takes \( n+1 \) inputs – the coefficients of the Bernstein polynomial – and selects one of them based on the output of the decoding block. [75] proves how this circuit correctly implements the degree-\( n \) Bernstein polynomial. This method can be used to implement many useful stochastic logic functions either directly, or via approximation. For example, [75] describes how this method can approximate many common mathematical functions (e.g., \( \sin, \tan, \exp \), etc.), while [77] describes a Bernstein polynomial that approximates the gamma correction function. These functions require higher-order polynomials than the simple circuit considered here (e.g., degree-5 for \( \sin \), degree-6 for \( \cos \)). The implementation of the degree-3 Bernstein polynomial is thus relatively simple, but serves to make a comparison between NML and CMOS implementations. Thus, Ch. [7.3.3] describes an NML implementation of the logic for a degree-3 Bernstein polynomial in order to compare to this approach.

7.1.2 Other Implementations

The approach posited by [75] has certain drawbacks, as this method can only operate on positive numbers (a bipolar representation to represent negative numbers is described in [33]), and many logic functions can only be approximated using Bernstein polynomials. For example, [75] considers degree-5 (to approximate \( \sin, \tan, \tanh \), etc.) and degree-6 (to approximate \( \cos, \exp \), etc.) Maclaurin polynomials. As the synthesis method to represent these functions in Bernstein polynomial form can only approximate these functions, an additional source of error is introduced (though [75] shows the total error is still less than for deterministic implementations for high error rates). [46] and [52] thus consider alternative stochastic implementations of various functions used in image processing (e.g., contrast stretching which uses \( \tanh \), KDE-based image segmentation which uses \( \exp \), etc.), and show how for complex functions, a stochastic implementation can offer significant improvements in terms of area, energy, and delay. As future work, an NML stochastic implementation should be compared to these methods, but the comparison to the approach in
Figure 7.2. Block diagram of circuit to implement a degree-3 Bernstein polynomial as described in [75].
serves as an initial exploration to show how NML might fit into a stochastic computing architecture.

7.2 Perpendicular Magnetic Logic

Before comparing to the method for implementing stochastic logic as presented in [75], stochastic architectures in the context of NML will be considered. Devices with perpendicular magnetization states – perpendicular magnetic logic (PML) – are considered. Recent progress with this device architecture is promising as many structures have been experimentally demonstrated, and properties unique to PML make it easier to layout circuits and lead to the deepest possible inherent pipeline.

PML embodies an alternative way to represent magnetic state to implement NML. Rather than representing binary state via the magnetization of *in-plane* nanomagnets, one could represent magnetic state via alternating stacks of cobalt/platinum in a *perpendicular* arrangement. The two approaches are depicted in Fig. 7.3. PML magnets can be defined in the same way as in-plane NML (henceforth called iNML to avoid confusion), or via focused ion beam irradiation (FIB) [40]. With FIB, the anisotropy of the magnets is destroyed in order to define devices, and can also be used to define the direction of dataflow through the devices. In the latter case, the region exposed to FIB has its anisotropy destroyed, resulting in magnetic reversal beginning at this point (since magnetic reversal begins at the point with the lowest anisotropy). It is therefore straightforward to ensure unidirectionality in PML without requiring a 3-phase (or 4-phase, etc.) clock for the magnets (as is required for iNML, as shown in Ch. 5.1.1). Another potential benefit of PML is described in [15], which showed that by defining the nucleation site with FIB, the required field strength to clock the devices varies significantly less (1-2%) than for iNML (up to ±10% as shown in [89]).

Single PML devices [9], inverter chains [29] and majority gates [11] have been demonstrated experimentally. More complex structures have been studied via simulation, e.g., a
systolic pattern matcher in [41] and a non-Boolean edge detector in [63]. The proposed

clock for PML devices is out of plane, and alternates sinusoidally [40]. Switching oc-
curs only when a magnet-pair is in a parallel configuration, and these metastable states

are driven through the circuits with each application of the clock pulse. Thus, the deepest

possible inherent pipeline is created – specifically, three magnets at minimum are required
to represent a single bit to ensure logically correct operation.

PML in theory lends itself well to operating in a stochastic computing paradigm in ei-

er a serial or parallel implementation. In a serial implementation, long streams of random

bits are fed through relatively simple processing elements. Simple processing elements al-

low one to take advantage of the strengths of PML (low energy, inherently pipelined opera-

tion) while mitigating the drawbacks from nearest-neighbor interaction and low switching
times compared to CMOS. A parallel implementation containing many copies of these

simple processing elements would similarly take advantage of the strengths of PML. The

benefits of PML when implementing simple, streaming processing elements has been ex-

pounded in the context of systolic architectures as in [63], which shows how PML can offer

improved performance compared to CMOS equivalents.

Specifically, there are two reasons why PML is considered instead of iNML for this

Figure 7.3. Comparison of iNML (left) to PML (right). [63].
study. First, in-plane signal crossing is required to realize many complex structures, including the majority-gate based adder adder in Ch. 5.4.1 (depicted in Fig. 5.3). In PML, signal crossing can be achieved by routing through a second magnetic layer as described in [30]. This method has been shown to work in simulation and in experiment – in contrast, signal crossing has been shown to work for iNML only in simulation (see Ch. 3.5). Note that the signal routed through the upper magnetic layer can be transmitted in one clock pulse, and that device coupling to the higher layer is in parallel (instead of anti-parallel as is usually the case between magnets in the same plane). Second, it is far simpler to ensure the correct number of magnets are present in inverter chains (i.e., wires) leading up to gates. For example, consider a multiplexer that implements the function \( Y = \overline{S} \cdot A + B \cdot S \). Since \( S \) and \( \overline{S} \) are needed, the critical path through the two required AND gates will be different (i.e., since \( S \) is inverted going into one AND gate, an additional magnet is needed in the critical path through that gate). Thus, the interconnect leading to the OR gate will require inverter chains of different length at its inputs. To account for this when employing iNML, one would need to increase inter-device spacing between one of the lines of magnets or use magnets of different sizes (for both of which it may be difficult to achieve fine-grained control). For PML, it is straightforward to control the size of the magnets. Though PML devices can be defined using liftoff, etc. as in iNML, FIB can also be used to define the size and shape of each device. Thus, if devices of varying sizes are needed in a circuit, this can be provided via FIB for a PML implementation (whereas for iNML it would need to be accomplished via traditional, variation-prone methods such as liftoff, etc.).

### 7.3 Benchmarking

Two studies are now presented that compare performance of CMOS and NML stochastic logic implementations. First, Ch. 7.3.1 discusses a comparison for a simple stochastic logic function (namely, multiplication) between CMOS (for deterministic and stochastic implementations) and PML (for a stochastic implementation). The CMOS stochastic
comparison includes projections based on the NRI benchmarking efforts [67], as well as HSPICE [81] simulations using the Predictive Technology Model (PTM) from Arizona State University (ASU) [7].

Next, Ch. 7.3.2 describes an implementation of the stochastic computing architecture in [75] for PML. This comparison is “on paper” versus simulation-based in order to show an initial exploration of potential mappings between stochastic computing architectures and the PML device architecture. To make the comparison to [75], area and delay are quantified for a PML implementation of the decoder/multiplexer block for a Bernstein polynomial of degree-3. Area is calculated based on the layout of the PML implementation, and delay based on the magnetic switching time multiplied by the critical path length. The expressions for area and delay from [75] are in terms of inverters and 2-input NOR, NAND, and OR gates. For the sake of simplicity, data from [67] is leveraged for 2-input NAND gates to determine expressions for PML to compare to the results in [75]. That is, the area of the PML layout (to be discussed in Ch. 7.3.3) is determined, and divided by the area of a 2-input NAND gate for 15 nm CMOS (thus, area for the PML implementation is quantified in terms of 15 nm CMOS 2-input NAND gates).

7.3.1 Simple Benchmark

Simple, circuit-level comparisons between PML and CMOS are now discussed. Specifically, a PML stochastic multiplier is compared to a CMOS stochastic multiplier and a CMOS deterministic multiplier. Described first are how area, energy, and delay were calculated for each, followed by a comparison of results.

7.3.1.1 CMOS Deterministic Multiplier

There are two methods for projecting the performance of a CMOS deterministic multiplier. First, [67] presents area, energy, and delay estimates for 1- and 32-bit adders. They also present estimates for a 2-input NAND gate, etc. These results are based on the 15 nm
technology node, and both high performance and low-power CMOS are considered. By following their methods, one can calculate the area, energy, and delay for an AND gate and a D flip-flop. The area of a gate simply depends on the number of p- and n-type transistors, and is calculated accordingly. Specifically, [67] calculates area by estimating the gate length and width (from layout diagrams) and multiplying by a “gate overhead factor” (this factor accounts for any extra empty space around the gates, and is derived empirically). For example, the area of an inverter is equal to \( 2p_m \times (3p_m + 2w_x)M_{gate} \) where \( p_m \) is the metal-1 pitch, \( w_x \) is the gate width, and \( M_{gate} \) is the gate overhead factor [67]. The same method is used in this work to calculate the area for an AND gate, D flip-flop, and 1-bit full adder. Note that for an AND gate (3 p-type and 3 n-type devices), the area for a NOR gate (2 p-type and 2 n-type devices) plus the area of one inverter (1 p-type and 1 n-type device) is assumed. The area of the flip-flop is calculated in a similar manner but assuming 10 p-type and 10 n-type devices. The adder calculation follows the assumptions in [67] (i.e., area for two XOR gates and three AND/OR gates, where the area of the XOR gate is estimated using NAND gates as reported in [67]).

Similar to the area calculations, [67] assumes energy and delay are a function of the number of transistors, as well as the intrinsic transistor and interconnect values (capacitance, voltages, etc.). Using the method and values reported in [67] for intrinsic transistor and interconnect energy and delay, the energy and delay for the 1-bit adder, AND gate and flip flop are calculated. Finally, with the area, energy, and delay results for a 32-bit adder, AND gate, and flip-flop, the performance of an N-bit array multiplier can be estimated. An array multiplier is simply composed of an array of full adders and requisite AND gates, and flip-flops at input/output are included to consider the cost of shifting data into the multiplier. Note that only low-power CMOS is considered, as in Ch. 5.4.1 as that corresponds with the likely application space for PML.

Energy and delay have also been projected for CMOS 32-bit ripple carry using ASU’s PTM [7]. The methods and results for gathering this data are presented in [43], and sum-
marized here for convenience. In [43], the 32-bit ripple carry adder was built from scratch with buffered transmission gate full adders. ASU’s PTM was used to simulate the circuit in HSPICE with inputs 0xFFFF + 0x0001. [43] measured the delay between when the input signal reached 50% of the maximum, and when the carry out reached 50% of the maximum, and measured energy by calculating charge dissipated during this time (static power was not included). Data was gathered for the 7, 10, 14, 16 and 20 nm technology nodes and for voltages between 0.3 and 0.9 V. Now, to calculate the multiplier energy and delay the 32-bit RCA data is augmented to account for the layout of an array multiplier. Thus, the energy/delay for the required AND gates, as well as for a single flip-flop at input and output, are added to the 32-bit RCA data. The delay for the AND gate and flip-flop were not included in [43], and thus were gathered via HSPICE simulation as a part of this work.

Finally, to calculate area for this approach, the estimate used in [43] is followed. Since the area of a device is limited by the area of the contacts, this approach estimates the area of the source, drain and gate contacts (it should thus be applicable to both non-FinFET and FinFET devices). Each contact is estimated to be one gate width wide by one gate width high (i.e., \(w^2\)). Here, the gate width is estimated to be \(w = 2 \times l\) (where \(l\) is the gate length specified by the PTM). The total area for a device is estimated by including three horizontal squares, one each for the area for the source, gate, and drain (each assumed to be equal to \(w^2\)). The source, gate, and drain thus have a width \(3w\) and height of \(w\), and area is therefore \(3w^2\). An additional spacing of \(0.5w\) is added around the perimeter, and so the total area is \(8w^2\) [43]. The total area for the PTM devices is therefore \(8w^2 = 32l^2\) for a single device. To estimate area for each circuit, this area is simply multiplied by the number of devices in the circuit (as defined in the HSPICE models).

7.3.1.2 CMOS Stochastic Multiplier

There are two methods used to calculate performance for the CMOS stochastic multiplier. As a CMOS stochastic multiplier is simply an AND gate, the first method is the same
as that for the deterministic multiplier in Ch. 7.3.1.1. That is, the data in [67] is leveraged to determine the area, energy, and delay of a 2-input AND gate, and a D flip-flop. Three D flip-flops are included: one for each input, and one for the output of the AND gate. The long shift operations required to match the precision of the deterministic approach are taken into account (recall that an input stream of length \( N \) will have precision of \( \frac{1}{2^N} \), so to compare to an 8-bit multiply a stochastic stream of 256 bits is required). In this work, it is assumed that the data streams through the stochastic multiplier (rather than, for example, accounting for 256 input/output shift registers).

In addition, results are presented for HSPICE simulations of a stochastic multiplier to calculate energy and delay. As in Ch. 7.3.1.1, HSPICE simulations are performed for an AND gate with two flip-flops at the input, and one at its output. The HSPICE simulations use ASU’s PTM as in [43]. Delay is calculated by measuring the time from 50% of maximum at input to 50% of maximum at output through each flip-flop and the AND gate. Energy is measured by calculating the charge dissipated during switching. Again, as in [43] static power dissipation is not taken into account. This likely makes projections somewhat optimistic, and so they should only be compared relative to each other, rather than taken as absolute energy consumption. The simulation is run for low-power CMOS (included in the ASU PTM), and for different supply voltages (0.3 to 0.9 V to match the multiplier data based on the adder projections in [43]). To calculate area for this approach, the estimate used in [43] is again followed.

7.3.1.3 PML Stochastic Multiplier

Estimates for the PML stochastic multiplier are also based on the data published in [67]. The switching time of the magnets is assumed to be 0.1 ns, and magnet width, height, and inter-device spacing are all assumed to be 15 nm (to match \( F = 15 \) nm in the CMOS projections in [67]). Area is calculated by assuming the total width of the structure is the
width of two D flip flops\(^\text{1}\) plus a majority-gate based AND gate, and the height is 5 magnets tall (to accommodate the majority gate). Each bit requires 3 magnets (the minimum that will allow for logically correct operation). Delay is calculated simply by the number of switching events associated with the AND gate, i.e.,

\[ N \times M_{\text{crit}} \times t_{\text{mag}} \]  

(7.3)

where \( N \) is the length of the stochastic stream (assumed to be 256), \( M_{\text{crit}} \) is the critical path through the majority-gate based AND (i.e., 3 magnets) and \( t_{\text{mag}} \) is the switching time of the magnets. Finally, energy is calculated by

\[ E_{\text{switch}} \times (5 \times N) \]

(7.4)

where \( E_{\text{switch}} \) is energy per magnet switching event as calculated per the method in \[67\] (which assumed magnetostrictive voltage-controlled clocking), and \( N \) is the length of the input stream. The energy is multiplied by 5*\( N \) to account for the number of magnets in the majority gate.

7.3.1.4 Comparison

Energy and delay results are plotted in Fig. 7.4 for all implementations: CMOS deterministic (NRI and PTM), CMOS stochastic (NRI and PTM), and PML stochastic (NRI). Recall that the CMOS NRI data (both deterministic and stochastic) is collected for the 15 nm node, while the PTM data is collected for the 16 nm and 7 nm nodes. For the PTM, voltages from 0.3 to 0.9 V are considered – note in the case of the CMOS deterministic data, 0.3 V was not a viable voltage (the circuit did not operate correctly) for the RCA in \[43\] and is thus not included. Recall that it is assumed that the input and output registers

\(^{1}\)Note that the traditional logic of a D flip flop is not implemented, but rather the inherent pipelining of PML is taken into account to latch the data.
are composed of only a single flip-flop, i.e., as if the data is streaming through the multiplier as if it’s part of another circuit (rather than accounting for long input/output shift registers). The conventional CMOS multiplier is assumed to be an 8-bit multiplier, and to match this the stochastic multipliers require a 256-bit stream to represent the data.

First, Fig. 7.4 shows how a PML stochastic multiplier might compete with a CMOS stochastic implementations. Delay for the PML implementation is lower when compared to the delay at the two lowest voltages for the CMOS-PTM estimates, and for the delay based on the NRI data. However, in all other cases delay is lower for the low-power CMOS. In terms of energy, the PML implementation is significantly lower than any of the estimates for stochastic CMOS.

Where stochastic implementations – and PML in particular – might offer the most improvements is in terms of area. To compare area, the estimates as in [67] are again considered for 15 nm CMOS and for PML (recall, magnet footprints are 15 nm²), and use the method in [43] to estimate area for the PTM data. Fig. 7.5 contains the area estimates for each of the three implementations. Notably, area is significantly lower for the stochastic case (which is expected, as it is simply an AND gate), and even lower for the PML implementation.

Finally, Fig. 7.6 presents the aggregate area-energy-delay product (AEDP) benchmark for each of the implementations to show how they compare when all three metrics are taken into account (note that for the CMOS-PTM implementations, only the lowest AEDP value for the range of voltages considered are shown). As expected, the significantly lower area and slightly decreased energy for the PML implementations lead to the lowest AEDP – at least 1 order of magnitude better (as compared to the conventional, 7 nm PTM projections) and up to 3 orders of magnitude better compared to the other cases. These results are encouraging, and further study is needed in the context of more complex circuits and applications to ascertain whether PML can provide similar improvements at those levels.
Figure 7.4. Energy and delay for deterministic (CMOS) and stochastic (CMOS, PML) multipliers. Only low-power CMOS is considered.

Figure 7.5. Area for deterministic (CMOS) and stochastic (CMOS, PML) multipliers. Only low-power CMOS is considered.
Figure 7.6. Area-energy-delay product for deterministic (CMOS) and stochastic (CMOS, PML) multipliers. Only low-power CMOS is considered, and only the lowest AEDP of the various voltages (0.3 to 0.9 V) is displayed for each PTM data point (for clarity).
7.3.2 Degree-3 Bernstein Polynomial Benchmark

Though the multiplier comparison is useful to show how PML might compete with CMOS deterministic and stochastic implementations, a more complex example could offer more insight. Therefore, relative performance is now compared between a PML implementation of the circuit in Fig. 7.2 to the results from [75]. First considered is how one might implement this circuit in PML, followed by a comparison to the deterministic and stochastic implementations using the results and methods in [75] (namely, estimates for area, delay, and area-delay product).

7.3.3 PML Implementation of Degree-3 Bernstein Polynomial

Described first is how the decoder/multiplexer circuit from [75] might be implemented using PML. For the sake of simplicity (since layout is done by hand), only the circuit to implement a degree-3 Bernstein polynomial is considered. Per [75], a full adder is required to implement the adder tree (i.e., to count the occurrence of 1s in the two independent stochastic input streams). The full adder can be implemented in PML with 3 majority gates as in Ch. 5.4.1. A 2-to-4 decoder takes the outputs of the full adder (SUM and COUT) as inputs to generate unique output signals that serve as the select signals for the multiplexer. The decoder can be implemented with 4 majority gates with one input fixed to implement AND functionality (as in [36], [40], etc.). Thus, the decoder generates the following outputs:

\[ S_0 = (SUM \cdot COUT) \]
\[ S_1 = (SUM \cdot COUT) \]
\[ S_2 = (SUM \cdot COUT) \]
\[ S_3 = (SUM \cdot COUT). \]

Finally, the 4-to-1 multiplexer (with select signals dictated by the decoder) can be implemented with 4 majority gates fixed to AND functionality, and three majority gates fixed to OR functionality (as in [36]). Thus, the multiplexer implements the function

\[ S_0 \cdot Z_0 + S_1 \cdot Z_1 + S_2 \cdot Z_2 + S_3 \cdot Z_3 + S_4 \cdot Z_4 \]

where each \( S_i \) is one of the select signals from the
2-to-4 decoder, and each $Z_i$ is one of the stochastic bit streams representing the coefficients of the Bernstein polynomial. Each of $Z_i$ is an input to the circuit (green squares marked $Z_i$ in Fig. 7.7), while each of $S_i$ is generated by the decoder and routed appropriately (each $S_i$ is indicated in Fig. 7.7).

The design is laid out by hand as depicted in Fig. 7.7. Dataflow directionality is accomplished via the FIB method described in Ch. 7.2 but FIB-irradiated regions are not indicated for ease of reading the figure. The direction of dataflow is indicated with inset arrows instead. Note that each majority gate is an inverting majority gate (per [36], [40]).

### 7.3.4 Comparison

The area of the circuit in Fig. 7.7 is $23N$ wide by $14N$ tall (where $N$ is the width of a square magnet in Fig. 7.7). The critical path length is 44 magnets long. With this data, the area and delay of the circuit can be estimated. Then, the data from [67] can be used in order to convert these results to an equivalent number of 2-input NAND gates. As with the projections for PML in Ch. 7.3.1.3, $N$ is assumed to be 15 nm, and spacing to be 15 nm, and the switching time of a single magnet to be 0.1 ns. Thus, area is $289,900 \text{ nm}^2$, and delay (critical path multiplied by magnet switching time) is 4.4 ns.

To make a direct comparison to [75], PML area is divided by the area of a 2-input NAND. The area of a 2-input NAND is calculated in [67], and is based on the number of p- and n-type devices. The area of the 2-input NAND is $360 F^2$ per [67], where $F$ is assumed to be 15 nm. Thus, area in terms of 2-input NAND gates is 2.98. In other words, the PML implementation in Fig. 7.7 takes up the same area as 2.98 15 nm CMOS 2-input NAND gates. Similarly, the delay of the PML implementation is divided by the delay of a 2-input NAND from [67], i.e., 0.1396 ns for low-power CMOS (i.e., supply voltage of 0.3 V). Thus, delay (i.e., critical path) in terms of 2-input NAND gates is 31.52. In other words, the delay through the PML implementation is equivalent to the delay through 31.52 2-input NAND gates. Therefore, area delay product (ADP) is 93.93 in terms of low-power CMOS 2-input
A summary of the ADP projections for the PML stochastic, CMOS deterministic, and CMOS stochastic implementations of a degree-3 Bernstein polynomial are presented in Table 7.1 for various values of $M$, where $M$ is the number of bits representing the data (and thus the stochastic stream has a resolution of $2^{-M}$ per [75]). For example, for a PML implementation with $M = 7$, ADP is calculated by multiplying 93.93 by $2^7$. The results in Table 7.1 show that in terms of ADP, the PML implementation is significantly lower than that of projections for both stochastic and deterministic CMOS implementations as presented in [75]. This agrees with the comparison of the multiplier in Ch. 7.3.1 which suggested that a PML implementation could achieve significant savings (mainly through reduced area). These results are encouraging, but future work is needed to compare to higher-degree Bernstein polynomials, and to make comparisons with data gathered in simulation (both for CMOS and PML). Finally, recall that the method of synthesizing stochas-
tic logic proposed in [75] has some drawbacks (e.g., it cannot operate on negative numbers, and can only estimate certain functions). As such, future work should also consider the image processing benchmarks in [46] and [52] (which is significantly more challenging due to the complexity of those functions).

7.4 Conclusions

I have shown how PML might fit into a stochastic computing architecture by making initial area, energy, and delay projections to compare to CMOS deterministic and stochastic implementations. For a simple multiplier, PML can potentially offer significant reductions in terms of area, moderate reduction in terms of energy, but results in longer latencies. Therefore, in terms of the AEDP, PML is estimated to be between 1-3 orders of magnitude lower compared to CMOS deterministic and stochastic implementations. For a degree-3 Bernstein polynomial, PML should be competitive with low-power CMOS in terms of the ADP. Further work is needed for the former benchmark in order to obtain data for the PML implementation that is based on micromagnetic simulation (rather than on “paper-based” calculations). Similarly, further work is needed for the latter benchmark to obtain projections based on simulation (both HSPICE simulation for the CMOS implementations, and micromagnetic simulation for the PML implementation).
TABLE 7.1

COMPARISON OF AREA DELAY PRODUCT (UNITLESS) FOR IMPLEMENTATIONS OF A DEGREE-3 BERNSTEIN POLYNOMIAL.

<table>
<thead>
<tr>
<th>M</th>
<th>CMOS Deterministic</th>
<th>CMOS Stochastic</th>
<th>PML Stochastic</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>99,207</td>
<td>28,160</td>
<td>12,023</td>
</tr>
<tr>
<td>8</td>
<td>152,745</td>
<td>56,320</td>
<td>24,046</td>
</tr>
<tr>
<td>9</td>
<td>222,615</td>
<td>112,640</td>
<td>48,092</td>
</tr>
<tr>
<td>10</td>
<td>310,977</td>
<td>225,280</td>
<td>96,184</td>
</tr>
<tr>
<td>11</td>
<td>419,991</td>
<td>450,560</td>
<td>192,368</td>
</tr>
</tbody>
</table>
NML is a promising candidate technology to replace CMOS for certain application spaces, specifically those that can exploit NML’s strengths (low energy dissipation in the magnets, non-volatility, etc.) while mitigating its weaknesses (relatively high delay, limitations imposed by nearest-neighbor interactions, etc.). This document described my contributions to the state of the art in NML research, and my efforts to identify the ideal application spaces for NML.

First, in Ch. 3 I described methods to bridge the gap from the device level to the circuit level. Prior to the onset of my research efforts, the majority of NML research had been performed at the device level. Methods were needed to understand device-to-device interaction in order to enable the design of more complex circuit operation. Therefore, in Ch. 3 I first presented a method to quantify the interactions between devices. This allowed me to show how to design more complex circuit structures (wire crossings and vertical interconnect/corner turns were used as examples). Another necessary component to enabling large-scale circuit operation is to understand how to move signals across clock wire boundaries. Therefore, in Ch. 3 I also presented a preliminary study of how to control clock groups with a multi-phase clock to transmit signals across clock wire boundaries.

My second major contribution, in Ch. 4 was an in-depth study of the effects of realistic clock wires on logical correctness and power consumption for NML circuits. I showed how the preliminary study in Ch. 3 had two limitations that needed to be addressed (only two clock groups were considered, and uniform fields were employed). The new study thus considered circuits consisting of three clock groups, as well as the realistic field distribu-
tions generated by our proposed, experimentally-demonstrated clock wires. I showed that increased clock fields (and thus currents in the clock wires) are required to ensure logically correct operation, resulting in increased power consumption. In this study I also showed how field interference should not exist in a single plane of clock wires, and that variation in the clock wire cladding should not have a significant effect on circuit operation and power consumption. Finally, I showed the effects of field interference when multiple planes of clock wires (and multiple planes of NML devices, which can provide increased density) were considered. Notably, field interference will exist but can be overcome by increasing currents in the clock wires – thus, a tradeoff exists between energy and density.

Next, in Ch. 5 I demonstrate how to ensure uni-directional dataflow in NML circuits through the use of multi-phase clocks in order to satisfy one of the five tenets for digital logic [86]. I also describe how to quantify circuit energy and delay in the context of the realistic, multi-phase clocks. Then, I quantify energy and delay trends for various designs (using inverter chains/local interconnect as an example). With these results, I project performance for an NML ripple carry adder and compare to low-power CMOS equivalents, showing that NML could offer up to a 96X improvement in terms of the energy-delay product. I also described materials-centric paths to achieve even lower energy consumption due to decreased clock field requirements. Specifically, I considered new magnetic materials with biaxial anisotropy, and explored the performance of these materials in the context of multi-phase clocks.

Next, in Ch. 6 I showed how NML could potentially be used to implement shift registers (which can be volatile, or non-volatile depending on the application). In the context of non-volatile shift registers, I showed how conflicts can destroy saved state even in the context of logically-correct, pipelined circuit operation. Next, I presented a design space study to identify shift register designs with high stability and low energy. The methods used in this study can guide future circuit design efforts, and I identified the designs that should be studied first in experiment. I also described two potential checkpointing applications.
for the shift registers and showed how NML shift registers might outperform the devices originally proposed for these applications.

Finally, in Ch. 7 I consider how PML (magnets exploiting perpendicular magnetic anisotropy to operate out of plane) might server in a stochastic computing architecture. Notably, I compare area, energy, and delay projections for a PML stochastic multiplier to CMOS deterministic and stochastic multipliers. The projections suggest that the main advantage of PML in a stochastic computing architecture may be in terms of area, as well as decreased energy, that lead to an AEDP improvement up to 3 orders of magnitude over low-power CMOS equivalents. I also considered a more complex circuit for which I designed a PML implementation, and compared to area and delay projections for CMOS equivalents. Results again suggest that the PML implementation could offer significant improvements over CMOS equivalents in terms of the area-delay product.

Through my research efforts I have helped advance the state of the art of NML circuit, application, and architecture design. Most importantly, I have shown that NML can offer improved performance over many equivalent CMOS applications and architectures. This is especially important as we continue to approach the fundamental scaling limits for CMOS and the need for novel computing paradigms increases. As much of my studies are based on projections for NML performance, we need future large-scale simulation efforts (e.g., for complex circuits like the ripple carry adder, stochastic circuits, etc.) as well as experimental efforts (e.g., to demonstrate correct operation of complex structures like wire crossings, to verify high stability/low energy shift register designs, etc.).


41. X. Ju, M. Niemier, M. Becherer, W. Porod, P. Lugli, and G. Csaba. Systolic pattern matching hardware with out-of-plane nanomagnet logic devices, 2013. ISSN 1536-125X.


71. S. S. P. Parkin, M. Hayashi, and L. Thomas. Supplementary online information for magnetic domain wall racetrack memory, 2008.


