Fabrication Variations and Defect Tolerance for Nanomagnet-based QCA

Michael Crocker, X. Sharon Hu, and Michael Niemier
Department of Computer Science and Engineering
University of Notre Dame
Notre Dame, IN 46556, USA
Email: {mcrocker,shu,mniemier}@nd.edu

Abstract

Tolerating defects and fabrication variations will be critical in any system made with devices that have nanometer feature sizes. This paper considers how fabrication variations and defects might lead to faulty behavior in Magnetic Quantum-dot Cellular Automata (MQCA) circuits and systems. Here, we leverage physical-level simulation to consider how fabrication variations might affect a circuit’s logical correctness. We then discuss how we can tolerate fabrication variations at the device, circuit, and architectural level.

1 Introduction

In this paper, we consider how fabrication variations will affect the logical correctness of circuit elements realized from a nanomagnet-based implementation of the Quantum-dot Cellular Automata (MQCA) device architecture. For MQCA, wires, gates, and inverters have all been realized, they operate at room temperature [11], and [5] estimates that if $10^{10}$ magnets switch $10^8$ times/second, they would only dissipate about 0.1 W of power. When the drive circuitry is included, [14] predicts that circuits could provide performance wins over state-of-the-art, low power CMOS when considering energy delay product. Devices can scale and remain non-volatile provided their size/shape remains above the superparamagnetic limit (approximately 20 nm). However, binary state in nanomagnets with feature sizes below the superparamagnetic limit can be stable for around 1 ms [16] – long enough to perform logical operations. Scaling can also decrease switching times [16].

Still, like any device with nanometer feature sizes, MQCA-based circuits could suffer from defect rates that are much higher than those for CMOS-based circuits. Thus, an MQCA-based circuit must not only perform better for some computational task of interest (to justify a technology transition), but realistically will need to do so with more faulty components. Fabrication processes envisioned for MQCA are similar to those for CMOS and fabrication variations should be similar as well. However, because MQCA devices process information in different ways than CMOS devices, defect tolerance mechanisms will be different. We study these issues here. We pay particular attention to how defect tolerance mechanisms affect performance – as an understanding of correctness/performance tradeoffs early in the device design process can facilitate experimental “mid-course corrections” to ensure that a new technology can ultimately best CMOS for a given application.

2 Background

Figs. 1a-b illustrate two important building blocks that would be used to construct MQCA circuits. A wire (Fig. 1a) is just a line of magnets that are antiferromagnetically coupled with each other. The basic logic gate in MQCA is based on the majority voting function. By setting one input of a majority gate to a logic ‘0’ or ‘1’, the gate will execute an AND or OR function respectively. In MQCA, the gate performs an inverting majority gate function (Fig. 1b). These structures have all been experimentally demonstrated at room temperature (see Fig. 1c,d [11]).
Figure 1. Schematic/experimental representations of (a)/(c) a wire segment and (b)/(d) a majority gate. Operating scheme of a wire: (e) (i) initial configuration, (ii) high-field (“null”) state, (iii) after the application of the input, and the final ordered state.

The structures illustrated in Fig. 1c,d were tested with a clock that took the form of a periodically oscillating external magnetic field that drove a system to an initial state, and then controlled the relaxation of the said system to a ground state. For example, Fig. 1e-i, illustrates a line (or “wire”) of nanomagnets that has relaxed to a logically correct, antiferromagnetically coupled ground state. In Fig. 1e-ii, the external field turns the magnetic moments of all magnets horizontally into a neutral logic state against the preferred magnetic anisotropy (i.e. along the hard axes of the magnets). This is an unstable state of the system, and when the field is removed, the nano-magnets relax into a new antiferromagnetically ordered ground state in accordance with the new input (Fig. 1e-iii.) [14] explored the use of copper wires wrapped by ferrite on the sides and bottom to provide local control of MQCA-based circuits. Nanomagnets would reside on the wire surface.

3 Process Variations and Defects

MQCA devices are simply three-dimensional blocks made from one layer of a magnetic material (e.g. permalloy or supermalloy). Conventional, imprint, or electron-beam lithography (EBL) are all candidate fabrication mechanisms. EBL was used to make the nanomagnets for the MQCA majority gate and wire experiments discussed in [11]. Although EBL is well accepted for patterning at the nanometer regime, it can introduce many variations that may effect the functionality of an MQCA-based circuit: (a) Over-exposure and under-exposure could lead to incorrect shapes. (b) Irregularities in the EBL gun can lead to “shot noise” that can cause rough edges. (c) The electrons that make up the beam are subject to scattering effects that can cause exposure of the resist outside of the desired shapes [10]. This can lead to bulging around the edges. (d) If the spacing between shapes is small, the resist in the gap may fail completely during development, leading to the merger of two or more shapes. The thin resist walls can also bend, creating nanomagnets with bulges and indentations [9]. (e) With liftoff, a pattern is made in resist and metal is evaporated over it. When the resist is dissolved, the metal that previously resided on the top surface is washed away, or lifted off. However, the deposited material can clump, and the liftoff process might remove a large chunk from the edges of the rounded rectangle shapes.

Scanning Electronic Microscopy (SEM) images of nanomagnet lines from [12] (see Fig. 2a) illustrate said variations. While it is difficult to get precise information about the entire 3D nanomagnet shape from this top-down view, the images indicate that the edges of the nanomagnets are missing magnetic material. The SEM image in Fig. 2b shows a horizontal wire next to a vertical wire. The magnetic material bulges between the two wires. Similar behavior has been observed for nanomagnets that are placed close together.

4 Explaining and Quantifying Faulty Behavior

Here, we consider how missing magnetic material, bulges and “shifts”, and edge roughness affect how well a magnet represents and propagates binary state. Important to this
Figure 2. (a) Experimentally realized magnets can have slanted edges. (b) This SEM image shows wires made using EBL. Proximity effects and other variations have led to bulging between the two wires. The lithography design called for 25nm between wires, but the fabricated wire are much closer. Also, note that the vertical wire bulges most significantly where it is adjacent to the horizontal wire.

discussion is how the magnetization (binary state) associated with a previous computation is “removed”. This is essential as it allows the nanomagnets that make up MQCA circuit elements to be re-evaluated with new inputs (see Fig. 1e). Fabrication variations can make this process more difficult – which can in turn lead to faulty behavior and undesirable logical states. We leverage micromagnetic simulations of supermalloy nanomagnets (the material used in [2]) performed with the OOMMF simulation suite [7] to study these issues. As in [6], we 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 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 $\left| \frac{dm}{dt} \right|$ dropped below a present number of degrees per nanosecond. We note that OOMMF is widely used and there is excellent correlation between simulation and experimental results (see [15, 13] for examples).

4.1 A Base Case

Before looking at defective systems, we first consider a simulation of a non-defective magnet to provide a basis for comparison – a 60x90x30 nm magnet (as in [14]), in isolation, where the y-component of magnetization was initially strongly negative (see bottom inset in Fig. 3(a)). We then applied an external magnetic field ($H_{\text{clock}}$) orthogonally to the magnet (i.e. along its hard axis) that continuously increased in magnitude to simulate the effects of a clock. A 15,000 A/m biasing field was simultaneously applied along the magnet’s easy axis. The biasing field is necessary as, after we null the magnet ($M_y = 0$), we want to “tip” it to the opposite polarization. Without a biasing field, the magnet would randomly tip up or down as the field applied along its hard axis was removed. This simulation determines the external field required to null this magnet given this local bias.

The “No Defects” curve in Fig. 3(a) illustrates how the y-component of magnetization changes as a function of the magnitude of the external field. As the magnitude of the external field ($H_{\text{clock}}$) increases, the magnitude of its y-component of magnetization decreases – eventually reaching 0 when the external field is approximately $0.5 \times 10^5$ A/m. The nanomagnet is now nulled and is tipped up by the biasing field. Note that if the magnitude of $H_{\text{clock}}$ continues to increase, it will work to keep the magnet nulled (see the “tail” of the “No Defects” curve). This explains why after the magnet initially changes polarizations, the y-component of its magnetization trends back toward 0 – the external field can be too strong and works to re-null the magnet. When $H_{\text{clock}}$ returns to 0, the y-component of magnetization is strongly positive. After the magnet has changed polarizations, we again applied an external field along the magnet’s hard axis and a biasing field of the same magnitude (but in the opposite direction) to return the magnet to its original state. The direction of the nulling field should not impact the magnitude of the field required to null the magnet. For a magnet with no fabrication variations, this is in fact the case as seen in Fig. 3(a).
4.2 Missing Magnetic Material

We then performed a similar simulation, but this time removed magnetic material from the top-right of one magnet—a possible fabrication variation as discussed in Sec. 3 and seen in Fig. 2. While the same biasing field was used in this simulation, the external field was applied in the same direction for both state transitions. These simulation results are also illustrated in Fig. 3(a). We consider the down-to-up transition first (see middle inset in Fig. 3(a)). Note that a stronger external field is required to null this magnet (approximately $0.6 \times 10^5$ A/m instead of $0.5 \times 10^5$ A/m). Magnetic moments tend to align along a magnet’s edge. In this simulation, the placement of the slant and the direction of the applied external field help to reinforce the initial downward polarization ($\downarrow$). For this same reason, the up-to-down transition (see top inset in Fig. 3(a)) can be accomplished when the magnitude of $H_{\text{clock}}$ is lower (approximately $0.4 \times 10^5$ A/m). (Only the first portion of this curve is shown—i.e. until the magnet is nulled—to improve graph readability.)

We now consider the above results in the context of a simple circuit—in this case a line of 7, 60x90x30 nm magnets spaced 16 nm apart (see inset at lower left in Fig. 3(b)). The line was initialized to a logically correct, antiferromagnetically coupled ground state and the y-component of magnetization in the first magnet was switched in order to simulate a new input to this line (as in Fig. 1e). $H_{\text{clock}}$ was then applied along the magnets’ hard axes in the positive x-direction and increased from 0 A/m to approximately 40,000 A/m in approximately 1,500 A/m increments\(^1\). The field was then allowed to relax back to 0 A/m in the same manner.

The results of three simulations are summarized in Fig. 3(b). We plot the magnitude of $H_{\text{clock}}$ versus $M_y$ for the 4th magnet in the line—as $M_y$ represents binary state. We first considered a wire with no defects. In this plot, one can see that $H_{\text{clock}}$ was approximately 20,000 A/m when the 4th magnet was nulled. When a misshapen magnet is introduced into the line (insets show magnets 3, 4, and 5), as before, the external field required to null the misshapen magnet can increase or decrease depending on the magnet’s initial state. As seen from the inset at the top left of Fig. 3(b), the field required to null the defective magnet is actually about 20% lower if magnets 3 and 4 are initially $\downarrow$ and $\uparrow$. However, if magnets 3 and 4 are initially $\uparrow$ and $\downarrow$, the magnitude of $H_{\text{clock}}$ must be about 14% higher.

\(^1\)40,000 A/m is the magnitude of $H_{\text{clock}}$ that a clock wire in [14] might produce assuming a current density of $10^7$ A/cm\(^2\) — and was considered to be to a practical upper limit. Realistically, placing nanomagnets closer to one another can increase coupling and lower the magnitude of $H_{\text{clock}}$ required to null the magnets.
in order to “reuse” this system to correctly evaluate a new input.

The above simulations were repeated with magnets that became progressively more misshapen. These results are summarized in Fig. 4(a). As fabrication variation gets worse, the misshapen magnet is in fact never nulled. While $H_{\text{clock}}$ begins to cause the 4th magnet’s $\gamma$-component of magnetization to trend toward 0, the missing magnetic material/slant helps to re-enforce its initial state. Moreover, as the amount of magnetic material missing from the edge of the 4th magnet increases, it becomes more difficult to null when $H_{\text{clock}}$ has a similar magnitude. After $H_{\text{clock}}$ reaches 40,000 A/m and relaxes back to 0 A/m, the 4th magnet settles back into its old (and now logically incorrect) polarization.

This suggests that we can leverage the external drive circuitry to help tolerate the effects of fabrication variation. Increasing the magnitude of $H_{\text{clock}}$ might increase the likelihood that the nanomagnets will behave in a logically correct manner. A stronger external field makes it easier to remove the state associated with a previous computation. However, more reliability comes at the expense of an increase in system energy. If $H_{\text{clock}}$ is never strong enough to remove the state associated with the previous computation, missing magnetic material can induce stuck at faults. That said, whether or not a wire (for example) is stuck at 0 or is stuck at 1 can depend on (a) where on the magnet material is missing from, and (b) what polarization state a group of magnets was initially in. For the example discussed above, if the misshapen magnet initially has a positive component of $M_y$, it can easily make a ↑ to ↓ transition. However, a ↓ to ↑ transition is more difficult suggesting that a “stuck-at-down” is most likely.

The simulations discussed thus far have assumed that $H_{\text{clock}}$ is applied uni-directionally. However, as noted in [14], $H_{\text{clock}}$ could also be sinusoidal. This could be advantageous from the perspective of minimizing overall system energy as the clock generation circuitry could be derived from the designs for adiabatic clocks [1, 17]. Moreover, as seen in Fig. 3(a), for a magnet with no fabrication variations, the magnitude of $H_{\text{clock}}$ (not the direction) is of most importance when working to null a given magnet. However, if the magnet is missing magnetic material, this is not necessarily the case.

In Fig. 4(b), we illustrate the results of 2 simulations designed to determine how an external parameter – the direction of $H_{\text{clock}}$ – might affect the logical correctness of some physical structure – a configuration of nanomagnets with fabrication variation. Like the simulation results summarized in Fig. 3(a), we assume a 60x90x30 nm magnet initialized such that its $\gamma$-component of magnetization was strongly negative. The first half of this simulation is essentially identical to the ↓ to ↑ transition captured in Fig. 3(a). However,
when considering the $\uparrow$ to $\downarrow$ transition here, the external field was this time applied from right-to-left (again with a bias of 15,000 A/m in a direction opposite to the magnet’s initial polarization). If $H_{\text{clock}}$ is applied in this manner, the magnitude of the external field must increase ($0.6 \times 10^5$ A/m vs. $0.4 \times 10^5$ A/m) to support the $\uparrow$ to $\downarrow$ transition.

These results will have to be carefully considered in MQCA circuit design. Consider a line of 3 nanomagnets where the middle magnet has missing magnetic material (i.e. as in Fig. 4(b)). If the initial state of this line is $\uparrow\downarrow\uparrow$ and $H_{\text{clock}}$ is applied from left-to-right, the third magnet will be more likely to end up in a “stuck at up” state per the discussion above. However, if the initial state of the line is $\downarrow\uparrow\downarrow$ and $H_{\text{clock}}$ is applied from right-to-left (with data still flowing from left-to-right), the third magnet is likely to end up in a “stuck at down” state. Essentially, a stuck at fault can occur – but it may change between stuck-at-up or stuck-at-down. The stuck at value is determined by the previous state of a group of magnets, the location of missing material and the direction of $H_{\text{clock}}$ – and can change based on what inputs are applied.

4.3 Bulging Nanomagnets

Above we looked at how missing edges might affect a magnet’s logical state and dataflow. We now consider what might happen if magnets “bulge.” Because bulging usually occurs where lithographically defined shapes are adjacent and close to each other, most nanomagnets in a wire will see similar bulging from neighbors on their left and right. To consider the effects of fabrication variations that manifest themselves as bulges, we simulated lines of 60x90x30 nm, 65x90x30 nm, and 70x90x30 nm magnets. For sufficiently strong nulling fields, both up and down inputs were successfully propagated down each of the aforementioned lines. However, it is more revealing to examine how bulges affect the quality of a signal’s binary state by considering an M-H plot analogous to Fig. 3(a).

Specifically, Fig. 5 plots the y-component of magnetization against an applied external field. (Again, a 15,000 A/m bias is used.) As one can see, the magnet with a smaller aspect ratio is actually easier to null. For example, the magnitude of $H_{\text{clock}}$ when $M_y = 0$ for the 65 x 90 nm magnet is lower than that for the 60 x 90 nm magnet. This would seemingly indicate that the lower aspect ratio magnets that result from bulging are an asset as the current density in a clock wire could decrease. While in some sense this is true, there are consequences when moving toward magnets with smaller aspect ratios.

When studying a line or “wire” of nanomagnets, as a magnet’s aspect ratio decreases, the magnitude of the y-component of magnetization for each individual magnet in the line decreases as well. Magnets tended to couple with one another in the x-direction more.
readily. From the perspective of binary logic, this represents a “weaker” 1 or 0. While one might look at Fig. 5a and note that when \( H_{\text{clock}} \) was 0, \( M_y \) was strong, this is only because the global bias field was applied continuously throughout this simulation. To better quantify the y-component of magnetization as a function of aspect ratio, we initialized a 60x90, 65x90, and 70x90 nm magnets in isolation such that \( M_y \) was initially strongly negative and then let each magnet relax to a ground state with no applied biasing field. We report the average Y component of their magnetization in the presence of zero bias (normalized to ±1) in Fig. 5b. We averaged \( M_y \) over all of the simulation mesh points and also averaged \( H_y \) 4 nm away from the magnet. As the aspect ratio decreases, the magnitude of the y-component of the field below the magnet decreases. This makes sense as the magnetization in the x-direction increases as aspect ratio decreases. Logically, as aspect ratio decreases, the y-component of magnetization decreases – resulting in a potentially “weaker” 0 or 1 and a “weaker” driver due to a lower biasing field. While further discussion is beyond the scope of this paper, a strong component of \( M_y \) is desirable as it can help facilitate other circuit structures (i.e. the driver to a crossover).

### 4.4 Shifted Nanomagnets

We now consider what might happen if the space between adjacent nanomagnets in a line is non-uniform. This will bring 2 nanomagnets closer together and move two others farther apart as seen in the insets of Fig. 6(b). (With no variations, each nanomagnet would be 16 nm away from the next.) In these simulations, we will consider \( M_y \) as a function of the x-component of the local magnetic field seen near the surface of a nanomagnet in a gap (see Fig. 6(a)). This measurement \( (H_{x-\text{local}}) \) is more insightful as it also incorporates the contribution of the magnetic material itself. One way to interpret this data is that in order to reuse a magnet to evaluate new inputs, this 60x90x30 nm magnet needs to “see” this value of \( H_{x-\text{local}} \). Depending on the size and shape of the surrounding magnetic material, a given value of \( H_{x-\text{local}} \) may be generated with a higher or lower external field.

We leverage \( H_{\text{local}} \) to consider how spacing variations between nanomagnets might affect dataflow in a 7 cell line with no missing magnetic material. Simulation results are illustrated in Fig. 6(b) where we plot \( M_y \) of the 5th magnet from the left versus the local field 8 nm to the left of the 5th magnet. We begin by examining the 16 nm curve – which is essentially a wire with no spacing variation between the magnets. Note that when the magnet is nulled,
$H_{local} = 2.8 \times 10^5 \text{ A/m}$. We next examine the curve for the simulation with a 20 nm spacing between the 4th and 5th magnets. Again, the 4th magnet is successfully nulled and that the wire functions correctly. However, while $H_{local}$ is nearly identical to the 16 nm case, the corresponding value of $H_{clock}$ when magnet 5 was nulled was greater than the magnitude of $H_{clock}$ for the 16 nm case – indicating that a stronger clock/more energy is needed to null magnet 5 due to weaker coupling with the adjacent magnetic material. Looking at the curves for 24 and 28 nm spacing between the 4th and 5th magnets, the magnitude of $H_{clock}$ is not sufficient to null all of the magnets in a line and a stuck at state ensues.

### 4.5 Edge Roughness and Magnetic Material

A magnet’s edge roughness and the magnetic material used to make it can also affect circuit functionality. While the details are beyond the scope of this paper, we do highlight two simulation results. First, as a magnet’s edge roughness increases, the magnitude of $H_{clock}$ required to null it generally increases as well (when compared to simulation results described by Figs. and 3(a) and 4(b)). Second, simulation results indicate that magnets made with a harder magnetic material (i.e. permalloy instead of supermalloy) may be more difficult to null, but also provide a stronger, local biasing field. Depending on the type of defect, sometimes a stronger local biasing field can allow fault free operation with a lower nulling field (as an individual magnet can provide more local control over nearby devices).

### 5 Defect Tolerance at the Architectural Level

Reconfigurable logic offers another way to avoid faults via architectural-level redundancy. For example, the PLA structure in [8] can be expanded to include more rows and columns such that defective crosspoints and/or interconnect can be avoided – increasing the probability that the desired set of logic functions can be mapped onto the faulty PLA. However, for MQCA, a larger PLA not only means a larger chip area, but also more/longer clocking wires to operate the control logic and interconnect associated with it. Therefore, redundancy in an MQCA PLA provides a way to trade power consumption for fault tolerance.

Consider the yield vs. fault rate study presented in [3]. This study indicates that a yield of 90% is possible given a fault rate of $10^{-3}$ and 10% redundancy. However, if the fault rate increases to $10^{-2}$, 400% redundancy is required. As seen in Sec. 4 increasing the magnitude of $H_{clock}$ provides another level of flexibility to circuit designers in terms of fault tolerance. According to [14], the magnitude of $H_{clock}$ increases linearly with the current density and linearly with the current in a given clock wire. However, the power increases with the square of the current. Therefore, doubling the magnitude of $H_{clock}$ implies that current in a wire will double while the power will quadruple. Thus, from the standpoint of performance and logical correctness, it is an interesting optimization problem to determine the most effective usage of the above mechanisms for fault tolerance. Together these techniques could allow for tolerance of a higher fault rate than can be achieved by either individually. However, one technique might be sufficient to provide the fault tolerance required to achieve a desired yield with the smallest increase in power.

To further quantify the tradeoffs discussed above, we worked with the ALU4 function from the Toronto 20 benchmark suite. Assume that the fault rate is normally $5 \times 10^{-4}$ with a clocking field of 20,000 A/m, and an increase in the clocking field to 24,000 A/m leads to a lower fault rate of $2.5 \times 10^{-4}$. We used the mapping tools proposed in [4] to map the ALU4 benchmark to the MQCA PLA. For the higher fault rate ($5 \times 10^{-4}$), the area needed to successfully map the ALU4 benchmark to the PLA with a yield of 96% is approximately 68,500 $\mu m^2$. For the lower fault rate ($2.5 \times 10^{-4}$), a PLA that has an area of only 24,500 $\mu m^2$ is sufficient to map the benchmark with 96% yield. These area estimates are based on 60x90nm magnets with 20nm spacings and wires crossings as described in [14].
To estimate the power consumption, the clocking wire design given in [14] is adopted, where the wire has a width of 2 $\mu$m$^2$ and a length of 4 $\mu$m$^2$ for an area of 8 $\mu$m$^2$. This means that the larger PLA will require 8,563 wires while the smaller PLA will require 3,063 wires. However, for the smaller PLA, the clocking field needs to be increased by 20% from 20,000 A/m to 24,000 A/m. This 20% increase in field strength means a 20% increase in current, and a 44% increase in power consumption. Therefore, if the power consumed by one clocking wire region is $P$, then the power consumption for the larger PLA with the lower clocking field is 8563$P$. The power consumption for the smaller PLA with the higher clocking field is 4411$P$. In this case, a 96% yield can be achieved with less power consumption if $H_{\text{clock}}$ is increased so that a smaller PLA can be used to map the logic. A systematic study of this tradeoff will be done in our future work.

6 Conclusions and Future Work

We have illustrated how fabrication variations might induce logical faults in MQCA-based circuits, identified faulty behavior unique to MQCA, and have demonstrated three different ways to mitigate the effects of faulty behavior – that can all be applied simultaneously to lower the overall fault rate. We will continue to leverage this work to determine what fabrication variations are most problematic in terms of both logical correctness and overall system performance to shape experimental work for this promising technology.

References