Reliable Nanowire Addressing via Randomized-Contact Decoders

Eric Rachlin and John Savage

Abstract

In January of 2007, researchers from Caltech and UCLA demonstrated a nanoscale memory with 10^{11} storage devices per cm [9]. This memory used a grid of nanoscale wires (NWs) to control bistable rotaxane molecules located at NW crosspoints. Similar crossbar-based storage technology, using larger feature sizes, was demonstrated by researchers at HP in 2004 [11]. In both cases the limiting factor on storage density was not wire pitch or device size, it was an inability to efficiently control individual NWs using lithographically produced mesoscale wires (MWs). To solve this, several types of NW decoders have been proposed. We model one type, the randomized-contact decoder, and provide strong bounds on its area overhead and defect tolerance.

1 Introduction

NW crossbars are a potential basis for both memories and circuits [5, 6]. Crossbar-based architectures would have densities a hundred times greater than today's CMOS. Unlike current VLSI, however, they would be assembled stochastically. This results in three fundamental challenges:

1. NWs are randomly assigned physical addresses
2. Testing is needed to configure crossbar control circuitry
3. Both permanent and transient faults must be tolerated.

As a means of addressing these challenges, we consider the **randomized-contact decoder** (RCD). A **nanowire decoder** is any device capable of addressing many individual NWs using MWs. The term RCD applies to any NW decoder in which MWs are coupled to NWs independently at random during decoder assembly. Like other proposed NW decoders, RCDs are assembled stochastically.

Our analysis shows that RCDs are efficient and robust. They reliably control a large number of NWs using many fewer MWs, even when NW/MW contacts fail. Our work also looks at the overhead of using an RCD. NWs in an RCD are randomly assigned physical addresses that must be discovered and stored. We compare two strategies for storing addresses and discuss a simple discovery procedure.

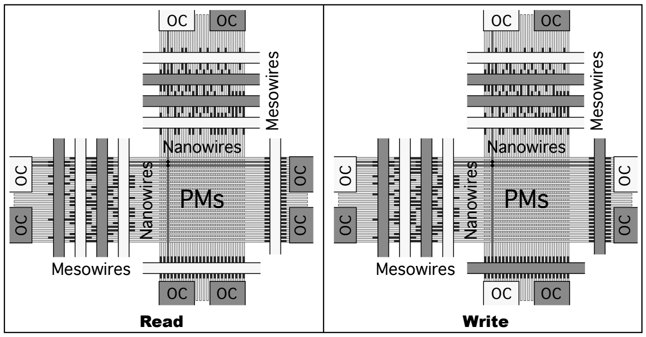

Figure 1: A crossbar-based memory in which OCs and MWs read and write data to PMs. In a read operation an OC at each end of a NW is disconnected from ground. Individual NWs are addressed by MWs causing current to flow only if the NWs' crosspoints conduct. In a write operation, OCs along each dimension apply a larger electric field, setting the conductivity of any crosspoint being addressed.

2 Crossbar-based Memories

Grids of long straight NWs can be produced using a variety of methods. Undifferentiated (i.e. identical) NWs can be stamped onto a chip, [3, 12]. Alternatively, many types of differentiated NWs can be grown off chip, collected in a large ensemble, then deposited fluidically [18]. Using either approach, a molecular layer can be deposited between two orthogonal layers of parallel NWs. This layer can consist of programmable molecular diodes (PMs) that switch between low and high resistances in a large electric field [2, 4]. A layer of amorphous silicon has also been proposed [8].

2.1 Nanowire Decoders

Once assembled, NW crossbars are controlled using a NW decoder along each dimension (see Figure 1). In the decoders, ohmic contacts (OCs) place voltages across groups of consecutive NWs while lithographically produced MWs gate (i.e. make nonconducting) subsets of NWs. The subset of MWs that gate a NW is called its **codeword**. Codewords are chosen randomly during decoder assembly.

A NW is addressed by activating one or both of its OCs along with all MWs that do not gate it. If the number of MWs is sufficiently large, the probability all other NWs get turned off is high. When NWs along each dimension of a crossbar are addressed, a large voltage can be used to set the conductivity of the PMs at their crosspoint. A smaller voltage can also be used to measure the conductivity of the crosspoint. These actions act as **write** and **read** operations, respectively.

For crossbar memories to access many individual crosspoints, NW decoders must control many individual NWs. The probability a decoder individually addresses many NWs depends on how it is manufactured and the number of MWs it employs.

To control undifferentiated NWs with MWs, two related decoders have been proposed. The first, a **masked-based decoder**, relies on randomly shifted lithographically-defined high-K dielectric regions interposed between NWs and MWs [1, 15] The second, an RCD, is analyzed here. One proposed method for making RCDs is to randomly deposit impurities, such as gold, between NWs and MWs [19]. Another option is to randomly place nanoscale high-K dielectric regions.

To control differentiated NWs, an **encoded NW decoder** has also been proposed. These decoders can be produced using modulation-doped NWs [6, 20] (NWs grown with patterns of lightly and heavily doped regions along their axis), or radially encoded NWs [17] (NWs with layers of removable shells). In both cases NWs with different encodings are grown separately, collected in a large ensemble, then randomly deposited on a chip. When connected to OCs NWs with each encoding can be separately addressed.

2.2 Modeling NW Decoders

To analyze the RCD we use a model introduced in [14]. Let M and g be the number of MWs and OCs, respectively, along each dimension of a crossbar. Let w be the number of NWs connected to each OC and $N = gw$.

Definition 2.1 *For $\alpha \gg 1$, a set, $\mathcal{S}$, of NWs is* **addressed** *if and only if a subset of MWs is activated such that:*

1. *Every NW not in $\mathcal{S}$ has a resistance that is at least α times that of every NW in $\mathcal{S}$*
2. *The resistance of all NWs not in $\mathcal{S}$, when combined in parallel, is at least α times that of the combined resistance of all NWs in $\mathcal{S}$.*

A NW, $\boldsymbol{n_i}$, is **individually addressable** *if and only if the set $\{\boldsymbol{n_i}\}$ can be addressed by some subset of MWs.*

For each NW $\boldsymbol{n_i}$, we represent its codeword with an M-bit vector, $\boldsymbol{c^i}$. When activating the j^{th} MW, let r^i_j denote the increase in $\boldsymbol{n_i}$'s resistance. Define the j^{th} bit of $\boldsymbol{c^i}$ as follows:

- $c^i_j = 0$ if $r^i_j \leq r_{low}$
- $c^i_j = 1$ if $r_{high} \leq r^i_j$
- $c^i_j = e$ if $r_{low} < r^i_j < r_{high}$, meaning c^i_j is in error.

where $r_{low} = cr_{base}$, $r_{high} = \alpha(N-1)(cM-c+1)r_{base}$, r_{base} is the maximum resistance of any NW when all MWs are off, and c is any positive integer.

Multiple types of NW decoder can be modeled using codewords. The distribution with which codewords are assigned to NWs depends on how a decoder is manufactured. In an RCD each bit of each codeword is an identically distributed independent random variable. For all c^i_j, let $p = Prob(c^i_j = 1)$, $q = Prob(c^i_j = 0)$ and $r = 1 - p - q$.

Before bounding the number of MWs required by an RCD, we must provide criteria that a decoder can satisfy to address many individual NWs. To build intuition, consider two NWs, $\boldsymbol{n_a}$ and $\boldsymbol{n_b}$, connected to a single OC. Suppose there is a k such that $c^a_k = 0$ and $c^b_k = 1$. By activating the k^{th} MW, $\boldsymbol{n_b}$ can be turned off while $\boldsymbol{n_a}$ remains on. If there is no such k, however, this may not be possible, in which case $\boldsymbol{n_a}$ may not be individually addressable. This line of reasoning is formalized as follows [14]:

Definition 2.2 *If for all k $c^a_k = 1$ when $c^b_k = 1$,* $\boldsymbol{c^b}$ **implies** $\boldsymbol{c^a}$. *If for all k $c^a_k \neq 0$ when $c^b_k = 1$,* $\boldsymbol{c^b}$ **possibly implies** $\boldsymbol{c^a}$.

Lemma 2.1 *In a NW decoder, NW $\boldsymbol{n_a}$ is individually addressable if no other NW possibly implies it. In an idealized error-free decoder, NW $\boldsymbol{n_a}$ is individually addressable if and only if no other NW implies it.*

To bound the number of MWs required by an RCD we can bound the number of MWs required for all or most NWs to have codewords not possibly implied by any other codeword.

3 Bounds on MWs

Let N_a be the number of individually addressable NWs in an RCD. Through simulation, Hogg *et al* [10] explore the conditions under which N_a is close to N. They demonstrate that as M passes a threshold near $4.8 \log_2 N$, the probability most NWs are addressable grows rapidly. Their empirical study, however, does not give bounds on M in terms of ϵ, the probability N_a is less than a desired value. It also fails to capture the impact of manufacturing errors.

We give tight bounds for both purposes using two different approaches. First we bound the expected value of N_a, $E[N_a]$, and use Hoeffding's Inequality [13] to choose M such that N_a is close to $E[N_a]$. Second we use the principle of inclusion-exclusion to choose M such that all NWs connected to an OC are individually addressable with high probability.

3.1 Bounds Using Expectation

From [16] we have the following bound on $E[N_a]$:

Lemma 3.1 *Let $N = gw$ be the total number of NWs contained in an RCD with g OCs, w NWs per OC, and M MWs. If N_a is the total number of individually addressable NWs,*

$$N(1 - (N-1)(1-pq)^M) \leq E[N_a] \leq N(1 - (1-pq)^M)$$

where $p = P(c_j^i = 1)$, $q = P(c_j^i = 0)$ and $p + q \leq 1$.

Let N_a^i be the number of individually addressable NWs connected to the i^{th} OC. $N_a = \sum_{i=1}^{g} N_a^i$, and since N_a^i are independent random variables ranging from 1 to w we apply Hoeffding's Inequality [13]. (We use 1 as the lower bound since even if no NW is individually addressable, all NWs can be addressed at once.) The inequality gives $P(E[N_a] - N_a \geq k) \leq e^{-2k^2/g(w-1)^2}$ and leads to the following theorem [16].

Theorem 3.1 *Let N_a be the total number of individually addressable NWs in an RCD with $N = gw$ NWs, g OCs, w NWs per OC and M MWs.*

$$P(N_a > \kappa N) \geq 1 - \epsilon$$

if $\kappa \leq 1 - \sqrt{-\ln \epsilon/(2g^*)} - (w-1)(1-pq)^M$ where $g^* = g(w/(w-1))^2$.

The following example uses this theorem to show 13 MWs address at least 1027 NWs with probability .99 when $g = 175$, $w = 8$, $N = 1400$ and $p = q = .5$.

Example 3.1 *Let $p = q = .5$, $g = 175$, $w = 8$, $N = 1400$, $\epsilon = .01$, and $\kappa = .733$. When $M = 13$, $\kappa = .733 \leq 1 - \sqrt{-\ln .01/(2*175*(8/7)^2)} - 7*(3/4)^{13}$. By Theorem 3.1 $P(N_a > 1026.2) = .99$*

When manufacturing defects cause codeword errors, $r = 1 - p - q > 0$. If r grows as g stays fixed, M must increase to keep κ constant. If $pq = .2$ (rather than $pq = .25$ in our error-free example) M must grow by a factor of $\ln(.75)/\ln(.8) = 1.29$. If $pq = .1$, the factor is $\ln(.75)/\ln(.9) = 2.73$. Even for relatively high defect rates M is not prohibitively large.

3.2 Bounds Using Inclusion/Exclusion

Again we use a bound from [16]:

Theorem 3.2 *In an RCD with M MWs, the minimum value of M such that all w NWs connected to an OC are individually addressable with probability $1 - \epsilon$ satisfies the following:*

$$\frac{\ln(w(w-1)/2\epsilon)}{-\ln(1-pq)} \leq M \leq \frac{\ln(w(w-1)/\epsilon)}{-\ln(1-pq)}$$

where the lower bound only holds if $M \geq (1 - pq)/(pq \min(p,q))$ when $\epsilon \leq .05$.

Consider an RCD with $w = 8$, $g = 133$, $p = q = .5$ and $M = 30$. We now show that with probability .99 at least 128 OCs are connected to all individually addressable NWs.

Example 3.2 *An OC is said to "fail" if it is not connected to all individually addressable NWs. When $w = 8$, $p = q = .5$ and $M = 30$, Theorem 3.2 states that OCs fail with probability at most $\epsilon = .01$ (since $30 > -\ln(5400)/\ln(.75)$).*

Since OCs fail independently, the probability that f or fewer OCs fail is bounded by the tail of a binomial distribution. The probability is at least $\sum_{i=0}^{f} \binom{g}{i} \epsilon^i (1-\epsilon)^{g-i}$. When $g = 133$ and $f = 5$, this sum is greater than .99.

As discussed in at the end of Section 3.1, manufacturing defects only increase M by a small amount.

3.3 Comparison with Other Decoders

In an encoded NW decoder each NW is randomly selected from one of C types. All w NWs connected to an OC are independently addressable if they all have different types. If $C \geq w(w-1)/(-2\ln(1-\epsilon))$, all NWs have a different type with probability at least $1 - \epsilon$ [7].

Depending on how MWs are encoded, the number of MWs required is between $\log_2(C)$ and C [17]. Since $\ln(1-\epsilon) \approx -\epsilon$ when ϵ is small, the lower bound $M \geq \log_2(C) \approx \ln(w(w-1)/2\epsilon)/\ln 2$ is within a small constant factor of the bound in Theorem 3.2. Furthermore, RCDs may be easier to manufacture since they do not require producing many types of differently encoded NWs.

In a masked-based decoder, randomly shifted lithographically defined regions determine which MWs control which NWs. Unless regions can be placed with sub-NW pitch accuracy, close to $M = 2(w-1)\ln(2(w-1)/\epsilon)$ MWs are required to individually address all w NWs connected to an OC with probability $1 - \epsilon$ [15]. This is significantly more MWs than required by either RCDs or encoded NW decoders.

4 Address Translation Circuitry

Each binary address in a crossbar memory corresponds to a different pair of orthogonal NWs. During operation, binary addresses are split into high and low order bits, then each set of bits is used to address a NW along one dimension of the crossbar. **Address translation circuitry** (ATC) along each dimension maps these bits to an OC and subset of MWs to activate. Since NW decoders are assembled stochastically, this mapping varies from decoder to decoder. The ATC needs programmable storage for the codeword of each NW it addresses. It must also associate an OC with each codewords.

The way in which binary addresses are mapped to codewords and OCs is called an **addressing strategy**. We use examples 3.1 and 3.2 to compare two different addressing strategies. In the first strategy, **Take What You Get**, each binary address is mapped to any arbitrary addressable NW. For each address, the ATC stores an M-bit codeword and $\log(g)$-bit OC address. In the second strategy, **All Wires Almost Always Addressable**, w consecutive binary addresses are all mapped to the same OC. An M-bit codeword is still stored for each address, but now the address' high order bits (along with knowledge of which OCs fail to have w addressable NWs) are used to associate OCs with codewords.

4.1 Area Estimates

For each addressing strategy we estimate the total area, A_T, required to produce a memory with approximately 2^{20} bits of storage. We use the approach of [7] and write:

$$A_T \approx 2\chi\beta + 2\lambda_{meso}^2 g \log_2 g + (\lambda_{meso} M + \lambda_{nano} N)^2$$

where λ_{meso} and λ_{nano} are MW and NW pitch respectively, χ is the area of a mesoscale memory cell, and β is the number of bits stored in each ATC. In our formula for A_T, $(\lambda_{meso} M + \lambda_{nano} N)^2$ is the area of the crossbar and MWs, $\chi\beta$ estimates the area of each ATC, and $\lambda_{meso}^2 g \log_2 g$ estimates the area of each demultiplexer used to activate OCs.

Since χ will be many times λ_{meso}^2 and we expect $\lambda_{meso} \geq 10\lambda_{nano}$, “Take What You Get” appears superior.

Take What You Get: In Example 3.1, $g = 175$, $w = 8$, $N = 1400$, $M = 13$, and $N_a \geq 1027$ with probability a least .99. If ATC stores a codeword and OC for each address, $\beta = N_a(\lceil \log g \rceil + M) = 21,567$ bits are needed. This gives:

$$A_T \approx 43,134\chi + 2608\lambda_{meso}^2 + (\lambda_{meso} 13 + \lambda_{nano} 1400)^2.$$

All Wires Almost Always Addressable: In Example 3.2, $g = 133$, $w = 8$, $N = 1064$, $M = 30$ and $N_a \geq 1024$ with probability at least .99. Again the ATC stores a codeword for each address. Rather then also store an OC, we store an “offset” to add to each address' high order bits. Since at most 5 OCs fail, the offsets require at most 3 bits per 8 addresses. The ATC uses $\beta = N_A M + 3 * 128 = 31,658$ bits, giving:

$$A_T \approx 62,208\chi + 1877\lambda_{meso}^2 + (\lambda_{meso} 30 + \lambda_{nano} 1064)^2$$

4.2 Codeword Discovery

To program the ATC, codewords must be discovered. An efficient discovery algorithms for encoded NW decoders is in [7]. An algorithm for RCDs appears in [10] with a slightly flawed analysis. A modified version works well in simulation [16], although the impact of codeword errors is unknown.

A simpler approach to codeword discovery is **exhaustive search**. Here the current between two OCs connecting w NWs is measured for all 2^M combinations of activated MWs. The 2^M measurements reveal the codewords of individually addressable NWs [16]. If measurements can be taken from g OCs simultaneously, the algorithm becomes very efficient for small values of M (such as 13 from Example 3.1) [10].

References

[1] Beckman, Johnston, Luo, Green, and Heath. Bridging dimensions: Demultiplexing ultrahigh-density nanowire circuits. *Science*, 310:465–468, 2005.

[2] Chen, Jung, Ohlberg, Li, Stewart, Jeppeson, Nielson, Stoddart, and Williams. Nanoscale molecular-switch crossbar circuits. *Nanotechnology*, 14:462–468, 2003.

[3] Chou, Krauss, and Renstrom. Imprint lithography with 25-nm resolution. *Science*, 272:85–87, 1996.

[4] Collier, Wong, Belohradský, Raymo, Stoddart, Kuekes, Williams, and Heath. Electronically configurable molecular-based logic gates. *Science*, 285:391–394, 1999.

[5] DeHon. Array-based architecture for FET-based, nanoscale electronics. *IEEE Trans. on Nanotechnology*, 2(1):23–32, 2003.

[6] DeHon, Lincoln, and Savage. Stochastic assembly of sublithographic nanoscale interfaces. *IEEE Trans. on Nanotechnology*, 2(3):165–174, 2003.

[7] Gojman, Rachlin, and Savage. Evaluation of design strategies for stochastically assembled nanoarray memories. *JETC*, 1(2):73–108, 2005.

[8] Gopalakrishnan, Shenoy, Rettner, King, Zhang, Kurdi, Bozano, Welser, Rothwell, Jurich, Sanchez, Hernandez, Rice, Risk, and Wickramasinghe. The micro to nano addressing block. In *Procs. IEDM*, December 2005.

[9] Green, Choi, Boukai, Bunimovich, Halperin, DeIonno, Luo, Sheriff, Xu, Shin, Tseng, Stoddart, and Heath. A 160-kilobit molecular electronic memory patterned at 10^{11} $bits/cm^2$. *Nature*, 445:414–417, 2007.

[10] Hogg, Chen, and Kuekes. Assembling nanoscale circuits with randomized connections. *IEEE Trans. Nanotechnology*, 5(2):110–122, 2006.

[11] Jung, Ganapathiappan, Ohlberg, Olynick, Chen, Tong, and Williams. Fabrication of a 34x34 crossbar structure at 50 nm half-pitch by UV-based nanoimprint lithography. *Nano Letters*, 4(7):1225–1229, 2004.

[12] Melosh, Boukai, Diana, Gerardot, Badolato, Petroff, and Heath. Ultrahigh-density nanowire lattices and circuits. *Science*, 300:112–115, 2003.

[13] Mitzenmacher and Upfal. *Probability and Computing*. Cambridge University Press, Cambridge, 2005.

[14] Rachlin and Savage. Nanowire addressing with randomized-contact decoders. In *Procs. ICCAD*, November 2006.

[15] Rachlin and Savage. Analysis of mask-based nanowire decoders, June 2007. Submitted to IEEE Trans. on Computers.

[16] Rachlin and Savage. Nanowire addressing with randomized-contact decoders, April 2007. Submitted to Theoretical Computer Science.

[17] Savage, Rachlin, DeHon, Lieber, and Wu. Radial addressing of nanowires. *JETC*, 2(2):129–154, 2006.

[18] Whang, Jin, and Lieber. Nanolithography using hierarchically assembled nanowire masks. *Nano Letters*, 3(7):951–954, 2003.

[19] Williams and Kuekes. Demultiplexer for a molecular wire crossbar network, US Patent Number 6,256,767, July 3 2001.

[20] Yang, Zhon, and Lieber. Encoding electronic properties by synthesis of axial modulation-doped silicon nanowires. *Science*, 310:1304–1307, 2005.