

Self assembled three-dimensional nonvolatile memories

Leon Abelmann,^{*} Niels Tas, Erwin Berenschot, and Miko Elwenspoek[†]

MESA⁺ and IMPACT Research Institutes

University of Twente, The Netherlands

(Dated: July 9, 2009)

The continuous increase in capacity of non-volatile data storage systems will lead to bit densities of one bit per atom by the year 2020. Beyond this point, capacity can be increased by moving into the third dimension. At the transition from two to three dimensions, we need to store data at $4 \cdot 10^{20}$ bit/m³, or bit cell volumes of (135 nm)³. The transition will require the development of new manufacturing technologies. We propose to use self-assembly of artificial sub-micron elements, either in loosely organised networks or perfect crystals. In the latter case, the cross-point architecture currently used in solid state memories can be extended to three dimensions. There are many storage principles which can be used. When using principles requiring electrical connection to the storage elements, we show the need for transistor based cross-talk isolation. In case of magnetic storage, cross-talk can be avoided by reusing the coincident current magnetic ring core memory architecture invented in 1953, which does not require electrical contact between the address lines. We show that writing into sub-micron magnetic cores is theoretically possible by means of the stray field generated by currents passing through the address lines. Calculations show however that the inductive readout used in old-fashioned ring core memories is not possible, but sufficient signal to noise can be obtained by implementing a frequency mixing technique. We illustrate that realisation of self-assembled three-dimensional ring core memory is in principle possible, by combining corner lithography and anisotropic etching into single crystal silicon.

^{*}Electronic address: l.abelmann@utwente.nl

[†]Freiburg Institute for Advanced Studies (FRIAS), University of Freiburg, Albertstr. 19, 79104 Freiburg, Germany

I. INTRODUCTION

The progress made in micro- and nanotechnology industry over the last six decades has relied heavily on thin film technology. The versatile process of thin film deposition, lithography and etching has provided society with powerful miniaturized computers, equipped with transducers, communication channels and storage capabilities. Nanotechnology will continue to provide us with more, more powerful and smaller systems, primarily because of the ever reducing size of components.

The continuing decrease in size inevitably leads to the emergence of fundamental limits. Since the tiniest units of information are found in non-volatile storage systems (such as the hard disk), the first encounter with those limits was found in that area, where it expressed itself in the thermal stability of written bits [Charap1997]. Non-volatile data storage will therefore move away from magnetism, possibly in the direction of ferro-electrics [Cho2005] or phase-change [Hamann2006]. Ultimately, we will store bits of information into single atoms [Bennewitz2002].

Thin film technology is inherently two-dimensional, so we are accustomed to measuring progress in the number of transistors or stored bits per unit area. From a user point of view however, only the number of elements per unit volume is of interest. It seems to be in place to design new strategies which continue on the path towards increasing volumetric, rather than surface densities. Since the fundamental limits express themselves first in non-volatile storage, the first steps into the third dimension will be taken in this field.

The key idea in this paper is to machine suitable nanoparticles and assemble them in a self-assembly process. Each nanoparticle will have functionality to store one or more bits of information. When assembled in three-dimensional structures, the volumetric data density can grow far beyond the limit of two-dimensional storage.

By equipping particles with electronic functionality, we open the route towards three dimensional processors. On the long term, we can envision mixtures of particles with embedded digital electronics, non-volatile memory and communication circuits, combined in cubic millimeter computers.

In this paper we will limit ourselves to non-volatile data storage however. We will discuss the data storage roadmap and estimate when limits will be reached in section II, where we argue that solutions based on today's three-dimensional storage systems are not capable to meet the density or price requirements. Addressing will be discussed in section III, where it is shown that a three plane addressing scheme is to be preferred for large arrays. Subsequently we give two examples, based

on transistors with memory elements (Section IV) and magnetic ring cores (Section V). Realisation of this three dimensional non-volatile memory will depend on our ability to steer self-assembly of nano-particles, which will be shortly discussed in section VI. Finally, we demonstrate in section VII how an actual three-dimensional ring-core memory can be assembled from a limited number of tetragonal unit cells.

II. THE CHALLENGE IN NON-VOLATILE DATA STORAGE

The world of non-volatile data storage can be divided into two realms. On one side we have mechanically addressed memories, where information is addressed by physical movement of a read/write head. Famous examples are the record player, generations of optical storage media (from CD to Blu-ray disk), magnetic tape and hard-disks. Since the bit dimensions are determined by the size of one single –or only a few– read/write heads and by positioning accuracies, the data density of these media is generally very high. Mechanical addressing is however relatively slow (ms), and usually applied in situations where large amounts of data need to be stored which are not needed frequently. The highest data density is found in magnetic hard disk storage, and lies currently around 800 Gb/in². The growth in hard disk data density fluctuates between 40 and 100% per year. Predictions are that magnetic storage is limited to data densities of about 10 Tb/in² [Thomson2007], which is another eight years from now at a growth rate of 40%. Probe storage technology [Pantazi2008] offers the possibility to increase the data density all the way up to one bit per atom data storage [Bennewitz2002]. At this point the data density will be in the order of 250 Tb/in², which will be in 2020 with an average growth rate of 60% per year. It is not clear however how areal data density will increase beyond this point.

On the other side, we have electrically addressed memories such as DRAM, FlashMemory, MRAM, Phase-Change RAM etc. Since every bit is connected to the outside world by means of wiring, data rates and access times are high (ns). The capacity, or better bit density, of these media is limited by lithography, and follows the semiconductor industry roadmap. If the half-pitch linewidth is Λ , the minimum cell size is $4\Lambda^2$, and with multilevel storage at n bits per cell, the bit density becomes $n/4\Lambda^2$. The International Technology Roadmap for Semiconductors predicts that in 2020 Λ will be 11 nm (www.itrs.net). At a foreseen 4 bits per cell, the data density in electrically addressed storage will be 5 Tb/in². Consequently, electrically addressed memories will be far too expensive to take over the role of mechanically addressed memories in large capacity data storage systems in which access times of milliseconds are not an issue.

Areal data densities only give information about the

relative cost of the system, but not about the final dimensions. From a user viewpoint, we are much more interested in capacity per unit volume. This is exactly the reason why tape recording systems are still persisting, even though their areal density is two orders of magnitude below the disk version of magnetic data storage. Assuming that in 2020 a probe recording system will fit into a package of about 1 mm thick, the volumetric data density at the atomic limit (250 Tb/in²) will be 4·10²⁰ bit/m³. This translates into a bit volume of (135 nm)³.

These volumetric data densities are very difficult to obtain with today's technology –let alone extend beyond that value–, mainly because the physical principle does not allow further decrease in bit size, or the manufacturing costs become too high. We give two examples, holographic storage and stacking of FlashMemory cells.

In holographic storage [Heanue1994] data densities are determined by the wavelength of light, which is currently at 405 nm in commercial systems [Chuang2008]. Even though laboratory demonstrations are given at 257 nm [Ramanujam2006], these wavelengths are too large for the (135 nm)³ entry point. It is extremely difficult to develop recording materials which are transparent at sub-100 nm wavelength, and therefore doubtful whether holographic storage can provide sufficiently high volumetric densities. Moreover, full 3D holographic systems will be large due to the size of lasers operating at those wavelengths and positioning systems required.

A second solution, which is already implemented today in Flash memory, is to thin down silicon wafers and stack solid state memories on top of each other. At 20 μm wafer thickness, the volumetric data density would be sufficient at the 11 nm node. Demonstrations of 75 μm thickness have already been given [Dellutri2006], and wafer thicknesses down to 5 μm seem possible [Morrow2006]. Even if the technological challenge of stacking tens of layers can be met, the solution would not be cost effective since stacking does not reduce the price per bit (50 individual dies are not more expensive than 50 stacked on top of each other).

Rather than stacking individual dies, one can also try to stack layers of memory elements on top of each other on the same base wafer [Tanaka2007]. One does save substrate material this way, but the process is not inherently cheaper, since again tens of the highest resolution lithography steps are required. (Again, passing a single wafer 50 times through the same expensive lithography line is not orders of magnitude cheaper as passing 50 individual wafers.) Moreover, the yield for each step has to be significantly higher as for stacking of individual wafers.

It appears to us that there is currently no viable solution for high capacity storage above the 10²⁰ bit/m³ limit, and radically new approaches need to be investigated.

III. THREE DIMENSIONAL CROSS-POINT STRUCTURES

To realize a three-dimensional storage system, the self assembled array of nano-particles needs to be addressed from the outside. One could use a random agglomerate of particles, all connected together, provided that the percolation is adequate (there are sufficient electrically connected paths through the network). By using proper encoding and data detection algorithms, information storage can be stored in an associative fashion [Hopfield1982, McEliece1987].

From an architectural point of view, it is however much simpler to start with crystals of particles, in which the position of every particle is exactly known by its three coordinates in the crystal. We can then simply use the cross-point architectures also used in two-dimensional solid state memories today (DRAM, Flash).

To address memory cells in a two-dimensional cross point architecture, we need to select two coordinates (bit and word lines, see figure 2a). As a result, $2n$ address lines need to be connected for a n^2 bit memory. For a three-dimensional cross-point architecture, the situation is more favourable, with $3n$ lines addressing n^3 bit, with the advantage getting bigger for larger memories. For example, a 8 Tbit memory would require about 3·10⁶ connection wires in 2D, and only 6·10⁴ in 3D.

For a 3D array, there are several options to address individual bit cells. A single core can be addressed by three individual wires (Figure 1a). This has the advantage that selected memory cells receive three signals, whereas the non-selected cells receive only one signal, which reduces false selection. This approach has the major disadvantage however that every address line has to be connected, so $3n^2$ connections have to be made to the array.

By selecting planes of wires, rather than individual wires, the number of connections reduces to $3n$ (figure 1b). In this case the bit cells have to be able to discriminate between receiving two and three signals, which is less favourable.

A compromise can be found by selecting cells by two planes of wires, and one single wire (figure 1c). In this case $2n+n^2$ lines are required. Writing information can be achieved by first selecting a line of bit cells by two planes of wires, and then one single cell by a single wire. In that case the majority of the cells only experience one signal, a single line of two signals and the selected cell three signals.

The properties of these three configurations are tabulated in table ???. For large arrays, the advantage of the low number of connections for option b becomes dominant, even though one loses selectivity. Therefore we will consider the three planes selection method in the remainder of this paper.

The three-dimensional cross-point array will be connected by means of electronic circuitry including drivers

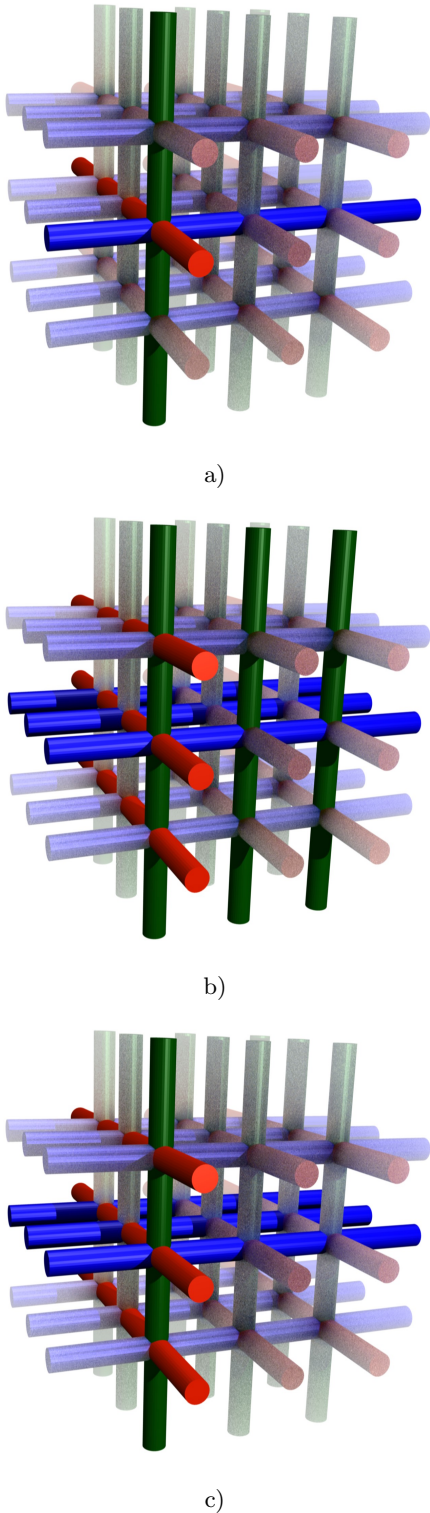


FIG. 1: Three addressing configurations requiring a) $3n^2$ b) $3n$ and c) $2n+n^2$ contacting wires. The selected wires are displayed in bright colour.

and multiplexers. In option b, the connection of these circuits to the array will be on the plane level, requiring (self-)assembly onto three orthogonal sets of n electrodes. The spacing of the electrodes is relatively large (downwards of 135 nm), and the number of electrodes is limited ($n=2 \cdot 10^4$ for a 8 Tb memory, which be smaller than $(3 \text{ mm})^3$). Electronically, this does not seem to be a challenging task.

Although individual cells can be addressed by selecting three planes, read-out can also be done in parallel for increased data rate. By selecting two planes, a line of bits is addressed simultaneously and n bits can be read out on the lines of the third plane. This reduces the required data-rate per bit cell considerably. For a 8 Gb/s readout speed of a 8 Tb memory for instance, planes have to be addressed at a rate of only 400 kHz.

IV. ELECTRICALLY ADDRESSED

For large numbers of bit cells addressed in the cross-point architecture, we need to consider cross-talk. In two-dimensional cross-point solid state memories, every bit is addressed by connecting a voltage source to its word- and bit line (figure 2a). The current passing through the connection can be used to write or detect information. Since all memory elements conduct current, parasitic paths exist which shunt current around the memory element under interrogation (green lines in figure 2a). These parasitic currents reduce signal-to noise ratio and strongly limit the array size. This can be avoided by integrating a diode within the memory cell. Since shunting currents pass in the opposite direction, parasitic currents are strongly suppressed. Even though the reverse currents through the diodes are very small, the combined effect through the many shunting paths can still be too large. In a 2×2 array for instance, there are only 6 parasitic paths, but this number increases rapidly with increasing array size. In general, the number of parasitic paths for a $n \times n$ array is A_n^{2n} , which can be calculated by iteration from

$$A_i^{2n} = \sum_{j=1}^{2n-1} A_{i-1}^j \quad (1)$$

$$A_2^{2n} = \left(\frac{1}{2}\right)^{2n} \sum_{j=1}^{2n-1} j \quad (2)$$

For large arrays the iteration can be approximated by $0.1 \cdot 2^{(n/2)}$ parallel paths. For a diode forward/reverse current ratio of 10^{11} , the maximum array size is limited to about 80×80 .

For larger arrays, we therefore need to move towards more aggressive rejection ratios, which can be achieved by using transistors rather than diodes (figure 2b). Since transistors are three-wire devices, an extra wire is connected, which - even though it can be shared between

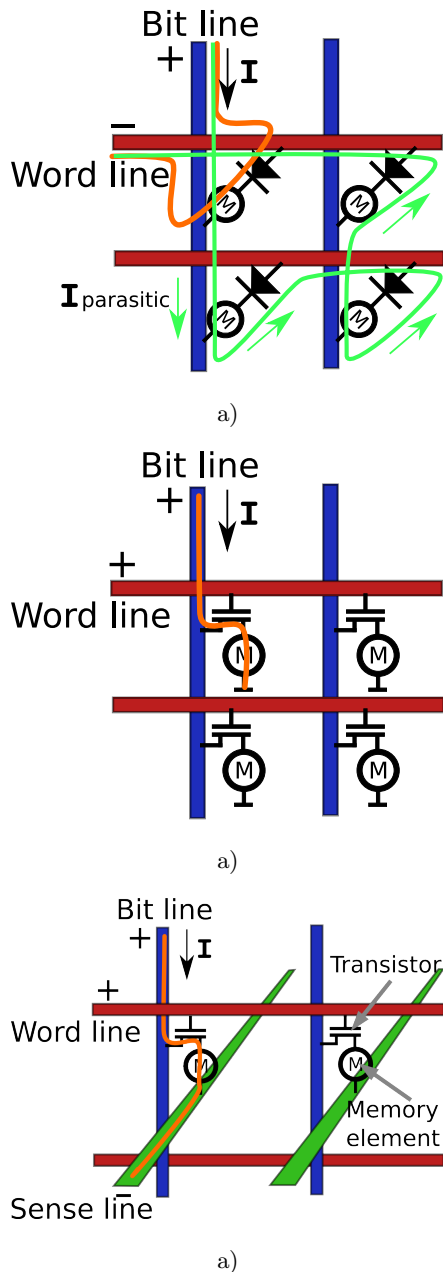


FIG. 2: (a) 2D cross-point architecture with diodes, (b) or with transistors. (c) 3D cross point architecture with transistors. M indicates the memory storage element, which can be based on various principles.

memory cells- complicates the architecture. Extending the cross-point architecture into three-dimensions however, there are three wires needed to address every bit, and the transistor geometry fits perfectly (figure 2c).

In order to realize transistor functionality in the nanoparticles, there is some advantage in using semiconductor base materials such as silicon. Sources and drains can be defined by doping opposite faces of particles, or by

assembly of n-doped and p-doped nanoparticles. Memory elements such as floating gates for charge storage (as in FlashMemory) or phase-change resistance wires exploiting the large difference in conductance between the amorphous and crystalline phase (as in PCRAM) can be incorporated at a third face or along edges of particles.

V. MAGNETICALLY ADDRESSED

Even though particles with active electronic functionality are extremely appealing, realisation of transistors inside nanoparticles or by assembly is technologically very challenging, since it requires excellent control over doping concentrations and interface quality. As a first step towards a three-dimensional storage system however, we can store data in a passive way, keeping the detection electronics outside the self-assembled structure. Therefore one needs to avoid the parasitic current paths.

A very elegant design of a three-dimensional cross point architecture without direct connection between the addressing wires was proposed by Jay W. Forrester [Forrester1951] in 1951, based on the invention of the magnetic core memory in 1950 by An Wang and Way-Dong Woo [Wang1950]. As Forrester writes in his introduction

In an ideal storage system it should be possible to arrange elementary storage cells in a compact three-dimensional array. Storage elements inside the volume should be selected by suitable controlling three coordinates along the edges of a solid array

At that time however, full 3D arrays proved to be difficult to manufacture, and a simpler stack of 2D planes was used in implementations.

In the magnetic ring core memory, data is stored in the direction of circulation of magnetisation of tiny magnetic rings (Figure 3 [1]). Information is written by means of the magnetic field of wires running through the center of the rings. Individual rings can be selected by passing currents only through wires selecting that ring (coincident current writing, invented by Forrester). Only in rings where two wires are actuated, the magnetic field is sufficiently high to induce switching. The rings therefore have to be able to discriminate between currents differing by a factor of two.

Depending on the direction of the current, clock or counterclockwise magnetisation patterns are written, representing “1” and “0” binary information. The use of magnetic ring cores was very elegant, since in such configuration the rotation of magnetisation is zero, and no magnetic stray fields emerge, therefore limiting the cross talk between elements.

In the 2D array, information was retrieved by attempting to overwrite information in the rings again and monitoring the change in flux. For this a current was applied to the two lines selecting a ring core (bit and word lines,

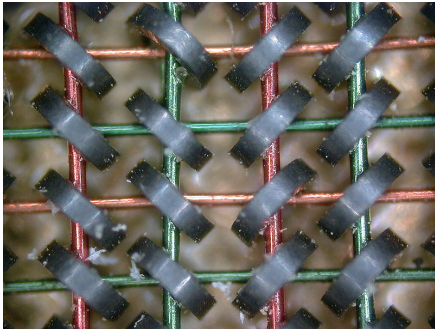


FIG. 3: Magnetic Core Memory

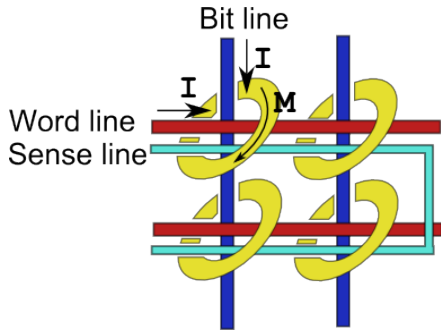


FIG. 4: Bit, word and sense lines are used to write and read information

figure 4). A third wire (sense line) passing through all cores was used to detect a change in magnetic flux, causing a induced voltage. From the presence or absence of this pulse the magnetisation direction could be derived. When a pulse was detected, an opposite current was applied to restore the state of the magnetic element (write-after-read scheme, invented by Wang).

Writing of sub-micron ring cores can be achieved in a similar way as in macroscopic ring cores. Since the magnetic volume is orders of magnitude smaller however, we need to consider the stability of the magnetisation direction at room temperature. This stability is determined by the energy barrier for magnetisation reversal, which should be higher than 40 kT for 10 year data retention. In appendix A we show that for ring diameters of 135 nm with wire diameters of 50 nm, the magnetic field is strong enough to overcome energy barriers of 120 kT.

Writing as in the macroscopic system therefore seems to be possible. The situation is different for readout however. When the magnetisation reverses, the flux through the circuit formed by the sense line changes, causing an *emf*. The flux change is proportional to the ring core wire diameter squared, so decreases rapidly with shrinking dimensions. The noise on the sense lines is proportional the square root of the wire resistance, so the signal to noise ratio decreases with a power 2.5. In appendix B we show that this prohibits this type of destructive readout

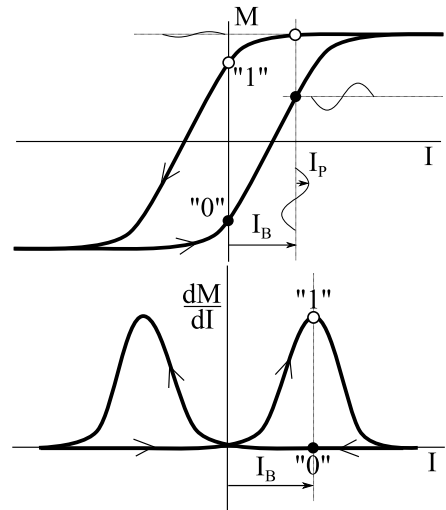


FIG. 5: High frequency modulation of the magnetisation can be used for non-destructive readout. a) Magnetisation curve b) Susceptibility

for nanometer scaled ring cores.

The main problem with destructive readout is that we get only one attempt to read out the information. Billing circumvents this problem by implementing a high frequency susceptibility measurement [Billing1966]. Rather than completely reversing the magnetisation in the core, a DC bias current I_B below the reversal threshold is passed through one of the addressing lines and small high frequency perturbation current I_P through another (figure 5a). Since the susceptibility (dM/dI) of the core depends on the direction of the current, the *emf* on the sense line can be used to determine the magnetic state of the ring (figure 5b). In appendix ?? we calculate that this method will allow for sufficient signal to noise ratio, but a trade-off needs to be made between wire diameter, modulation frequency and data rate.

VI. SELF ASSEMBLY OF ARTIFICIAL ELEMENTS

The relatively large dimension of 135 nm for bits required for a volumetric density of $4 \cdot 10^{20}$ bit/m³ suggest the use of nanotechnology. Fabrication of nanoparticles with those dimensions will be relatively cheap in 2020. Already today, methods exist to assemble these particles afterwards in large three-dimensional structures, similar to the way nature uses atoms and molecules to assembly larger structures.

The general concept is to fabricate large amounts of identical particles (spheres, cubes, tetrahedra, etc.) which are treated in such a way that they spontaneously self-assemble in liquid or gaseous environments. One method is to produce particles which have hydrophobic and hydrophilic surfaces. When immersed in water, the

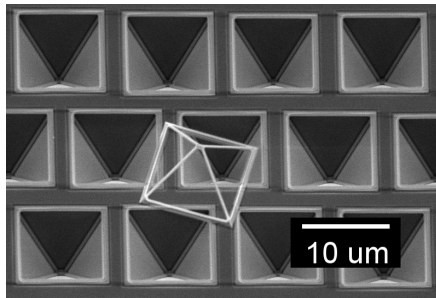


FIG. 6: Wireframe produced by corner lithography

hydrophobic surfaces bind together, producing irregular or crystalline agglomerates. Glotzer [Glotzer2007] gives an excellent overview of the state of the art in this field, listing many different types of particles and realized assemblies.

Up to now, the particles are passive, which is useful if one wants to realize fluidic filters [Seo2007] or photonic crystals [Fleming2002]. Although not yet demonstrated, several feasible paths to nanoparticles with internal degrees of freedom, or with electronic functions, with designed sterical and chemical anisotropy, are being considered - it is only a matter of time until they become available. Nanoparticles with a form commensurate with the crystal structure of silicon can be made by exploiting the great dependency of the etch rate in some high pH solutions. Adding an electronic circuit on the particles is not a major step. The technology allows for modification of the vertices, edges and faces of the tetrahedrons by means of corner lithography [Berenschot2008, Sarajlic2005e]. Figure 6 shows an initial result of a free SiN wireframe produced by this method. The size of the particle is determined by the etched inverse structure in the Si, which can be seen below the particle. The dimensions are determined by lithography, and can be reduced below 100 nm. Since the particles are embedded in before release, it is possible to deposit different materials in vertices, edges and faces.

Certainly nanoparticles can be modified to steer self-assembly into crystals so that regular structures similar to the cross-point structures shown in figure 1 can be grown. Parts of the nano-particles can for instance be made hydrophobic so that the assembly of amphiphilic particles into crystal structures similar to self-assembled structure in nature will become possible, see for instance [Rycenga2008, Murphy2005, Park2004b]. Alternatively, parts of the particle can be covered with organic agents or biological that get charged in a solution; by tuning the polarity quasi ionic crystals can be grown (see for instance [Shevchenko2006]). Even base-pair matching in DNA strands is demonstrated as a method for self-assembly [Rothmund2006]. After assembly, connections can be fixed by using for instance polymerized adhesives [Gu2004]. For an active assembly, some of the con-

nections between particles need to be conductive, which can for instance be achieved by diffusion of gold contacts [Gu2007]. For larger scale particles (300 μm), integration of electronics inside substrates has already been successfully demonstrated [Saedi2008].

Self-assembly of anisotropic nano-particles is currently in its infancy, however the emerging analogy with the growth of molecular crystals is quite encouraging. Regardless of the technique being used, it should in principle be possible to realize smarticles with transistor functionality or magnetic ring cores.

VII. SELF ASSEMBLED RING CORE MEMORY

In the last part of this paper we will explore whether it would *in principle* be possible to realise a 3D magnetic ring core memory by means of self assembly of a limited number of substructures produced by corner lithography. Two functions need to be assembled, a three dimensional array of wires (the word, bit and sense lines) and magnetic ring cores. There are many different ways in which a three dimensional array of wires can be self assembled, starting from tetrahedra, cubes or pyramids for instance. Table ?? lists the properties of unit cells constructed from these three basic elements. To realise 3D cross point architectures, for each of those basic elements eight elements are needed to realise a unit cell (Figure ??). Tetrahedra are the simplest elements, with exactly the right number of corner points. Starting from cubes, only two types of elements are needed (with and without wires). For tetrahedra and pyramids, four different elements are combined into the unit cell.

Also the magnetic ring core has to be defined by vertices. An efficient method is using tetrahedra. Figure ?? shows a unit cell of eight tetrahedra, with addressing lines at the corners of a square magnetic ring. Of course, the addressing lines have to be located within the magnetic ring for inductive read-out, and should be electrically isolated from it. Placing the wires close to the magnetic ring has the advantage of an increase in field strength. The magnetic elements do not have to touch, since stray field coupling will lead to flux closure structures.

This particular solution requires 5 different types of elements. There are other solutions, with more favorable positioning of the magnetic core, but at this point it is sufficient to realise that in principle it is possible to construct a 3D magnetic ring core memory from a limited number of simple building blocks. Figure ?? shows an impression of such an architecture, where the extension of the addressing lines can be clearly seen.

The technological challenges to realise a structure like this are formidable: We can make silicon tetrahedrons in a size range from a few tens of nanometers to a few micrometers, and we know ways to add electrical conductors of various metals along the edges. But these complex structures needs several types of tetrahedra. They have to be assembled in the right way, which is not trivial at all.

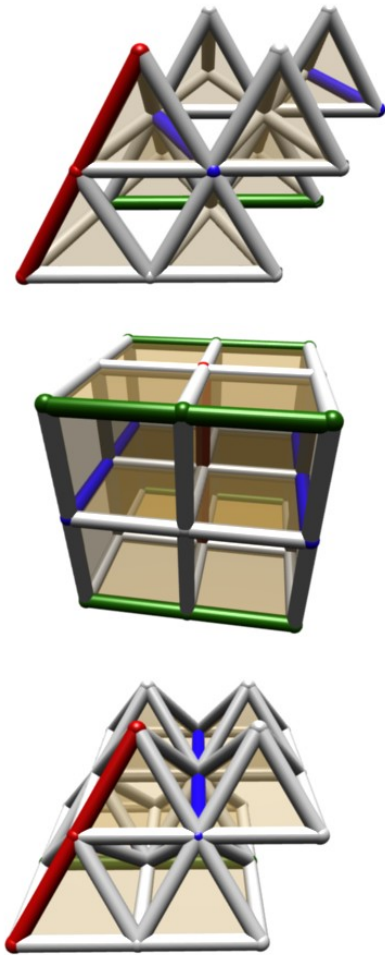


FIG. 7: Self assembly of wire frames (tetrahedra, cubes and pyramids) into units cells which can assemble into a 3D cross-point addressing architecture. Each unit cell contains eight smarticles.

Futhermore they should conduct electrical current along lines passing a vast number of particles, whilst keeping the impedance low.

	Option a	Option b	Option c
configuration	3 wires	3 planes	2 planes, 1 wire
number of connections	$3n^2$	$3n$	$2n+n^2$
selectivity	$1/3$	$2/3$	$2/3$

TABLE I: Comparison of the three addressing configurations, tabulating the number of connections to the array and the required selectivity.

	Corners	Vertices	Unit cells	Unique elements
Tetrahedron	4	6	8	4
Cube	8	12	8	2
Pyramid	5	8	8	4

TABLE II: Units cells for 3D cross point architectures, using tetrahedra, cubes or pyramids as building blocks.

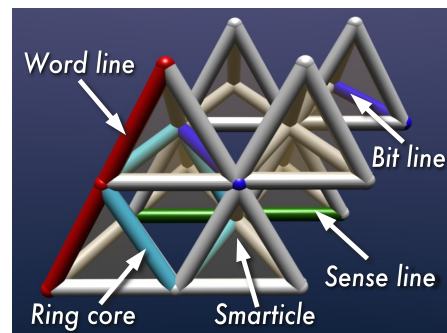


FIG. 8: By assembly of smarticles with different materials in the edges, a unit cell for a ring-core bit can be realized.

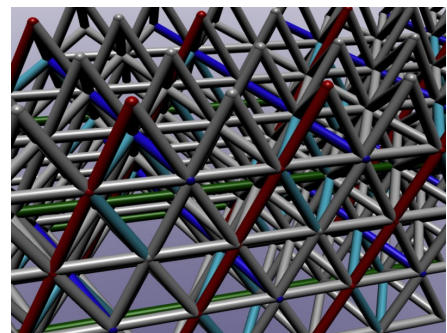


FIG. 9: Repetitive assembly of unit cells lead to a 3D ring core memory structure.

Acknowledgments

This work was supported by the Excellence Initiative of the German Federal and State Governments.

VIII. CONCLUSION

We expect the increase in data density for non-volatile, two-dimensional data storage to come to a halt when the bit spacing reaches atomic dimensions, which will be at a data density of about $4 \cdot 10^{20}$ bit/m². To continue the capacity of storage systems, we need to move from surface to volumetric storage. Using typical values for package sizes, one can estimate that the onset of three-dimensional data storage will have to start at a volumetric density of $4 \cdot 10^{12}$ bit/m³, or bit cell volumes of (135 nm)³.

By adopting the cross-point architecture from today's solid state memories, and extending it into three dimensions, a simple addressing scheme can be made. For large arrays, selection by three orthogonal planes is favourable since it reduces the number of connections to the outside.

If the memory cells are electrically addressed, transistors are required to reduce parasitic currents. Since transistors are three-terminal devices, they fit very nicely with the three-dimensional cross point architecture. As memory elements, phase change wires could be integrated.

Parasitic currents can be avoided by re-using the magnetic addressing strategies used in ring core memories from the early years of computer technology. We calculated that storing data into magnetic ring cores at sub-micron dimensions is possible, using coincident current writing at current densities well below the electromigration threshold, and energy barriers largely sufficient for 10 years of data retention. The small size of the rings prevent destructive readout at sufficient signal-to-noise ratio, but a susceptibility technique, combined with frequency mixing can certainly be applied.

A promising strategy for the realisation of three-dimensional memories could be the self assembly of artificial sub-micron elements (smarticles). Such elements can be realised by combining edge-lithography techniques and anisotropic etching. The first experiments into this direction are encouraging.

We show that a ring core memory can in principle be (self-)assembled from eight tetragonal wire frame elements per memory cell, consisting of five different types of tetrahedra each having different combinations of conducting, non-conducting and magnetic edges. Other configurations are however possible.

We therefore believe that the concept of a self-assembled three-dimensional ring core memory is technically feasible, and will provide an intriguing first step towards a solution for data storage beyond the atomic limit.

Appendix A: Write strategy

The stability of the magnetisation direction in a magnetic ringcore is determined by the energy barrier for magnetisation reversal, which should be higher than 40 kT for 10 year data retention. The maximum available energy for writing, ΔE , equals BMV , where B is the magnetic field from the conductor, M is the magnetisation of the ring material (about 1 MA/m) and V its volume. Starting from the (135 nm)³ bit cell design point, we assume a magnetic ring with a diameter D of 135 nm and wire diameter $2r$ of 50 nm. For simplicity, at the axis of the ring a conductor is positioned with the same diameter. The magnetic field B circulating the conductor is proportional to the current, which is in principle limited by electron migration current density j_{max} (approximately 10^{11} A/m²).

$$B = \mu_0 \frac{j_{max} r^2}{D} [T] \quad (A1)$$

The maximum current for a 50 nm diameter wire becomes 200 μ A, which limits the maximum magnetic field at 68 nm distance to 582 μ T. With the volume of the ring equal to $2\pi^2 Rr^2$, the energy barrier

$$\Delta E = \frac{\mu_0 \pi^2}{kT} j_{max} M r^4 [J] \quad (A2)$$

is in the order of $5 \cdot 10^{-19}$ J or about 120 kT, which is largely sufficient for long term non-volatile storage. Note that the energy barrier is independent on the ring diameter D , but strongly depends on the wire radius r . So in principle storing data in magnetic ring cores at these small dimensions is possible.

Appendix B: Read strategy

When the magnetisation reverses in the ring core, the flux through the circuit formed by the sense line changes, causing an *emf*. The flux ϕ [Vs] is proportional to the cross section of the magnetic rings

$$\phi = \mu_0 M \pi r^2, \quad (B1)$$

where r is the radius of the core wire (not to be confused with the core radius $D/2$). As the dimensions shrink, the flux reduces by a square power relation. Reversal can be detected if the *emf* caused by the change in flux is larger than the background noise on the sense wires. For bandwidths in the order of 10 Mb/s, the noise

for inductive readout is dominated by DC resistance of the sense wire [Klaassen1992], $S_N(\omega) = 4kTR$ [V²/Hz]. In order to estimate the signal-to-noise ratio (SNR) of the readback signal, we do not actually need know the pulse shape of the *emf* with time, but we can follow a matched filter approach [Koren1991].

Assuming an arbitrary pulse shape $f(t) = d\phi/dt$, Fourier transformed to $F(\omega)$, the energy in the pulse, relative to the noise, equals

$$E_w = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{|F(\omega)|^2}{S_N(\omega)} d\omega \quad (B2)$$

Koren states that one can always find a matching filter (MF), to shape the pulse in the optimum shape, and E_w can be optimized to SNR_{MF}^2 .

The chance on a read-back error, given the effective SNR_{eff} ratio equals

$$P_{err} = \frac{1}{2} \text{erfc} \left(\frac{\text{SNR}_{\text{eff}}}{2\sqrt{2}} \right) = \frac{1}{2\pi} \int_{\text{SNR}_{\text{eff}}/2}^{\infty} \exp(-t^2/2) dt \quad (B3)$$

For a raw error rate (without coding) of 10^{-4} , the SNR should be larger than 7.4 (8.7 dB) for a simple peak detection scheme.

If the pulses are well separated in time, and the the matched filter is applied $\text{SNR}_{\text{eff}} = \text{SNR}_{MF} = \sqrt{E_w}$. Pulse width and bandwidth are related, and cannot be made arbitrarily small. Assuming for instance a Lorenzian pulse shape (after matching) with pulse width a [s] and height f_{max} [V]

$$f(t) = \frac{f_{max}}{1 + \left(\frac{2t}{a}\right)^2} \quad (B4)$$

with f_{max} related to ϕ by

$$\int_{-\infty}^{+\infty} f(t) dt = \frac{1}{2} \pi f_{max} a = \int_{-\infty}^{+\infty} \frac{d\phi}{dt} dt = 2\phi \quad (B5)$$

$$(B6)$$

we find that

$$F(\omega) = 2\phi e^{\pi a |\omega|} \quad (B7)$$

and using (B2) therefore

$$\text{SNR}_{MF} = \sqrt{E_w} = \frac{\phi}{\sqrt{4kTR\pi a}} \quad (B8)$$

which should exceed 7.4. Assuming a 25 nm radius sense wire with a length L of 100 μ m and a specific resistance ρ of $1.7 \cdot 10^{-8}$ Ω m, R becomes 866 Ω . In that case a should be smaller than about 10^{-15} s. Micromagnetic simulations show that these short switching times

cannot be reached, certainly not under low field conditions [Hollinger2003]. The situation worsens when dimensions shrink. If we use a design length scale Λ , than the total flux ϕ scales with the cross-section of the magnetic ring, so Λ^2 . The resistance R of the sense wire scales with length over area, so $1/\Lambda$. The SNR therefore scales with $\Lambda^{2.5}$, which is not very benign. In the era of ring core memories, the core diameters were in the order of $300 \mu\text{m}$ [Werner1967]. We are now aiming below $1 \mu\text{m}$, and therefore loose over six orders of magnitude in signal to noise ratio. Destructive readout is therefore a dead end road and an alternative read-out technique needs to be investigated.

Appendix C: High frequency susceptibility readout

Rather than completely reversing the magnetisation in the core, a DC bias current I_B below the reversal threshold can be passed through one of the addressing lines and small high frequency perturbation current I_P through another (figure 5a). Since the susceptibility (dM/dI) of the core depends on the direction of the current, the *emf* on the sense line can be used to determine the magnetic state of the ring (figure 5b). Assuming a fraction of the magnetisation fluctuation γ of about 10%, the signal on the sense line will have an amplitude of $f_{max} = \gamma\phi\omega_p$. The noise power is now determined by the bandwidth of the detector B and equals $4kTRB$. The SNR of the readback signal can thus be calculated to be

$$\text{SNR} = \frac{\gamma^2 \phi^2 \omega_p^2}{4kTRB} \quad (\text{C1})$$

from which we can calculate that the minimum sense frequency

$$\omega_p > \frac{\sqrt{\frac{4}{\pi^3} kT \rho L B \text{SNR}}}{\gamma \mu_0 M r^3} \quad (\text{C2})$$

is a strong function of the wire diameter (which is set equal for the ring and sense line). For parameters identical as above ($r=25 \text{ nm}$), a sense frequency over 40 GHz is required for a bandwidth of 1 MHz (data rate of 500 kb/s), which is an engineering challenge. A start can be made however with a sense frequency of 1 GHz and a bandwidth of 1 MHz, which will require a wire radius of about 90 nm.

The high frequency fields are present throughout the array, and will cause induced voltages on sense lines of non-selected rings. Billing uses a very elegant technique to combat this cross-talk, in which the biasing current is also modulated at a different frequency. So

$$I_B = I_{BS} + I_{B0} \sin(\omega_B t) \quad (\text{C3})$$

$$I_P = I_{P0} \sin(\omega_P t) \quad (\text{C4})$$

By properly selecting the offset bias current I_{BS} , a non-linear region in the hysteresis loop can be selected. Assuming for simplicity

$$M(I) = -M_0 + \alpha I^2 \quad (\text{C5})$$

$$I = I_{B0} \sin(\omega_B t) + I_{P0} \sin(\omega_P t) \quad (\text{C6})$$

$$f(t) = \frac{\partial \phi}{\partial t} = \mu_0 \pi r^2 \alpha \frac{dI^2}{dt} \quad (\text{C7})$$

we find that the output signal has a response at double, sum and difference frequencies:

$$f(t) = \mu_0 \pi r^2 \alpha \left[\omega_B I_{B0}^2 \sin(2\omega_B t) + \omega_P I_{P0}^2 \sin(2\omega_P t) - (\omega_B - \omega_P) I_{B0} I_{P0} \sin((\omega_B - \omega_P)t) + (\omega_B + \omega_P) I_{B0} I_{P0} \sin((\omega_B + \omega_P)t) \right] \quad (\text{C8})$$

The sum and difference frequencies are not present for rings which are not selected, since for those rings either I_B or I_P is zero. By using a lock-in technique on $(\omega_B - \omega_P)$, cross talk can therefore be reduced. Taking a bandwidth in the order of MHz, and sense frequencies in the order of GHz, the reduction is at least 60 dB. An additional advantage of this method is that the lock-in technique uses frequency as well as phase information, which increases the SNR at the same time.

- [1] Image courtesy of H.J. Sommer III, Professor of Mechanical Engineering, Penn State University. 5
- [Bennewitz2002] Bennewitz, R., Crain, J. N., Kirakosian, A., Lin, J. L., McChesney, J. L., Petrovykh, D. Y., and Himpsel, F. J. (2002). Atomic scale memory at a silicon surface. *Nanotechnology*, 13:499–502. 2
- [Berenschot2008] Berenschot, E., Tas, N., Jansen, H., and Elwenspoek, M. (2008). 3d-nanomachining using corner lithography. pages 729–732, In *3rd IEEE International Conference on Nano/Micro Engineered and Molecular Systems, NEMS 2008*, pages 729–732, Sanya. 7
- [Billing1966] Billing, H., Rüdiger, A., and Schilling, R. (1966). Coincident current magnetic film memories using locked hard-direction films having microshape anisotropy. *Journal of Applied Physics*, 37(3):1367–1373. 6
- [Charap1997] Charap, S. H., Lu, P. L., and He, Y. (1997). Thermal stability of recorded information at high densities. *IEEE Trans. Magn.*, 33(1):978–983. 2
- [Cho2005] Cho, Y., Hashimoto, S., Odagawa, N., Tanaka, K., and Hiranaga, Y. (2005). Realization of 10 tbit/in.² memory density and subnanosecond domain switching time in ferroelectric data storage. *Applied Physics Letters*, 87(23):232907 – 1. 2
- [Chuang2008] Chuang, E., Sissom, B., Harris, R., Malang, K., Bergman, C., Hill, A., Bell, B., and Curtis, K. (2008). Demonstration of holographic read-only-memory mastering, replication, and playback with a compact reader. *Japanese Journal of Applied Physics, Part 1: Regular Papers and Short Notes and Review Papers*, 47(7 PART 2):5909 – 5911. 3
- [Dellutri2006] Dellutri, M., Pulici, P., Guarnaccia, D., Stopino, P., Vanalli, G., Lessio, T., Vassallo, F., Di Stefano, R., Labriola, G., Tenerello, A., Lo Iacono, F., and Campardo, G. (2006). 1 gb stacked solution of multilevel nor flash memory packaged in a lfbga 8 mm by 10 mm by 1.4 mm of thickness. *Thermal, Mechanical and Multi-Physics Simulation and Experiments in Micro-Electronics and Micro-Systems. Proceedings of EuroSimE 2006 (IEEE Cat. No.06EX1341)*, pages 5 pp. –, pages 5 pp. –, Piscataway, NJ, USA. 3
- [Elwenspoek2009] Elwenspoek, M. and Berenschot, E. (2009). Single crystalline silicon, chemically anisotropic Nanotrahedrons. *To be published*.
- [Fleming2002] Fleming, J., Lin, S., El-Kady, I., Biswas, R., and Ho, K. (2002). All-metallic three-dimensional photonic crystal with a large infrared bandgap. *Nature*, 417(6884):52 – 55. 7
- [Forrester1951] Forrester, J. W. (1951). Digital information storage in three dimensions using magnetic cores. *J. Appl. Phys.*, 22(1):44–48. 5
- [Glotzer2007] Glotzer, S. C. and Solomon, M. J. (2007). Anisotropy of building blocks and their assembly into complex structures. *Nature Materials*, 6(8):557 – 562. 7
- [Gu2004] Gu, Z., Chen, Y., and Gracias, D. (2004). Surface tension driven self-assembly of bundles and networks of 200 nm diameter rods using a polymerizable adhesive. *Langmuir*, 20(26):11308–11311. 7
- [Gu2007] Gu, Z., Ye, H., Bernfeld, A., Livi, K., and Gracias, D. (2007). Three-dimensional electrically interconnected nanowire networks formed by diffusion bonding. *Langmuir*, 23(3):979–982. 7
- [Hamann2006] Hamann, H. F., O’Boyle, M., Martin, Y. C., Rooks, M., and Wickramasinghe, H. K. (2006). Ultra-high-density phase-change storage and memory. *Nature Materials*, 5(5):383 – 387. 2
- [Heanue1994] Heanue, J. F., Bashaw, M. C., and Hesselink, L. (1994). Volume holographic storage and retrieval of digital data. *Science*, 265(5173):749 – 752. 3
- [Hollinger2003] Hollinger, R., Killinger, A., and Krey, U. (2003). Statics and fast dynamics of nanomagnets with vortex structure. *Journal of Magnetism and Magnetic Materials*, 261(1-2):178 – 189.
- [Hopfield1982] Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proc Natl Acad Sci U S A*, 79(8):2554–2558. 3
- [Klaassen1992] Klaassen, K. B. and van Peppen, J. C. L. (1992). Noise in thin-film inductive heads. *IEEE Transactions on Magnetics*, 28(5 pt 2):2097 – 2099. 10
- [Koren1991] Koren, N. L. (1991). Matched filter limits and code performance in digital magnetic recording. *IEEE Transactions on Magnetics*, 27(6 pt 1):4594 – 4599. 10
- [McEliece1987] McEliece, R. J., Posner, E. C., Rodemich, E. R., and Venkatesh, S. S. (1987). Capacity of the Hopfield associative memory. *IEEE Trans Inf Theory*, IT-33(4):461–482. 3
- [Morrow2006] Morrow, P. R., Park, C.-M., Ramanathan, S., Kobrinsky, M. J., and Harmes, M. (2006). Three-dimensional wafer stacking via cu - cu bonding integrated with 65-nm strained-si/low-k cmos technology. *IEEE Electron Device Letters*, 27(5):335 – 337. 3
- [Murphy2005] Murphy, C. J., Sau, T. K., Gole, A. M., Orendorff, C. J., Gao, J., Gou, L., Hunyadi, S. E., and Li, T. (2005). Anisotropic metal nanoparticles: Synthesis, assembly, and optical applications. *Journal of Physical Chemistry B*, 109(29):13857 – 13870. 7
- [Pantazi2008] Pantazi, A., Sebastian, A., Antonakopoulos, T. A., Bachtold, P., Bonaccio, A. R., Bonan, J., Cherubini, G., Despont, M., DiPietro, R. A., Drechsler, U., Durlg, U., Gotsmann, B., Haberle, W., Hagleitner, C., Hedrick, J. L., Jubin, D., Knoll, A., Lantz, M. A., Pentarakis, J., Pozidis, H., Pratt, R. C., Rothuizen, H. E., Stutz, R., Varsamou, M., Weismann, D., and Eleftheriou, E. (2008). Probe-based ultrahigh-density storage technology. *IBM J. Res. Dev.*, 52(4-5):493–511. 2
- [Park2004b] Park, S., Lim, J.-H., Chung, S.-W., and Mirkin, C. (2004). Self-assembly of mesoscopic metal-polymer amphiphiles. *Science*, 303(5656):348–351. 7
- [Ramanujam2006] Ramanujam, P., Lohse, B., and Berg, R. (2006). Organic materials for uv holographic and digital storage. *Proceedings of the SPIE - The International Soci-*

- ety for Optical Engineering*, 6252:62520 – 1, volume 6252, pages 62520 – 1, USA. 3
- [Rothmund2006] Rothmund, P. (2006). Folding dna to create nanoscale shapes and patterns. *Nature*, 440(7082):297–302. 7
- [Rycenga2008] Rycenga, M., McLellan, J. M., and Xia, Y. (2008). Controlling the assembly of silver nanocubes through selective functionalization of their faces. *Advanced Materials*, 20(12):2416 – 2420. 7
- [Saeedi2008] Saeedi, E., Kim, S., and Parviz, B. (2008). Self-assembled crystalline semiconductor optoelectronics on glass and plastic. *J Micromech Microengineering*, 18(7):–. 7
- [Sarajlic2005e] Sarajlic, E., Berenschot, J. W., Krijnen, G. J. M., and Elwenspoek, M. C. (2005). Fabrication of 3d nanowire frames by conventional micromachining technology. 1:27–30, In *Proceedings of the 13th International Conference on Solid-State Sensors, Actuators and Microsystems 2005, Seoul*, volume 1, pages 27–30, Los Alamitos. IEEE Computer Society Press. 7
- [Seo2007] Seo, Y. H. (2007). Nanomesh fluidic filter using self-assembly of colloidal nanospheres and surface tension. *Applied Physics Letters*, 90(12):123514 –. 7
- [Shevchenko2006] Shevchenko, E. V., Talapin, D. V., Kotov, N. A., O’Brien, S., and Murray, C. B. (2006). Structural diversity in binary nanoparticle superlattices. *Nature*, 439(7072):55 – 59. 7
- [Tanaka2007] Tanaka, H., Kido, M., Yahashi, K., Oomura, M., Katsumata, R., Kito, M., Fukuzumi, Y., Sato, M., Nagata, Y., Matsuoka, Y., Iwata, Y., Aochi, H., and Nitayama, A. (2007). Bit cost scalable technology with and plug process for ultra high density flash memory. *Digest of Technical Papers - Symposium on VLSI Technology*, pages 14 – 15, pages 14 – 15, Piscataway, NJ 08855-1331, United States. 3
- [Thomson2007] Thomson, T., Abelman, L., and Groenland, J. P. J. (2007). Magnetic data storage: past, present and future. In Azzerboni, B., Asti, G., Pareti, L., and Ghidini, M., editors, *Magnetic Nanostructures in Modern Technology*, NATO Science for Peace and Security Series Sub-series: NATO Science for Peace and Security Series B: Physics and Biophysics, pages 237–306. Springer Verlag, Berlin. 2
- [Wang1950] Wang, A. and Woo, W. D. (1950). Static magnetic storage and delay line. *Journal of Applied Physics*, 21(1):49–54. 5
- [Werner1967] Werner, G., Whalen, R., Lockhart, N., and Flaker, R. (1967). A 110-nanosecond ferrite core memory. *IBM J. Res. & Dev.*, pages 153–161.

