Will nanotechnology change IT paradigms?

G Aeppli, P Warburton and C Renner

There is a clear imperative to continue the IT revolution, even in the face of limits both from the laws of physics, as well as the challenge of managing complexity. Beyond changing the nature of hardware itself, nanotechnology and the replacement of classical logic gates with devices exploiting quantum mechanics could drive new paradigms for computer operating systems, going beyond Turing/von Neumann which is essentially the basis for all of modern IT.

1. The IT revolution

The information technology (IT) revolution is one of the defining features of our age, and promises the delivery of ever more sophisticated and personalised services at progressively lower cost. The drivers for the success of modern IT have been:

- exponentially increasing transistor density in integrated circuitry,
- bit density in magnetic storage media,
- communications channel capacity (wired, optical, and wireless) — all of which advanced simultaneously due to progress in microfabrication, micro-magnetics, and optics.

Remarkably, the underlying elements, most notably the transistor, used for switching, and the magnetic domain, the fundamental unit for data storage, have not changed over this period of tremendous growth. However, over the coming decade or so as these paradigms are pushed into the nanometre regime, fundamental limits will be reached:

- based on current trends, where ever fewer electrons are switched in each operation, we will soon need to switch less than one electron per operation, something which is at odds with what is physically possible (see Fig 1),
- as the magnetic domains representing bits are reduced, they will lose their stability relative to the thermal fluctuations present at room temperature,
- ever thinner gate oxides will result in larger leakage currents, and ever larger power consumption.

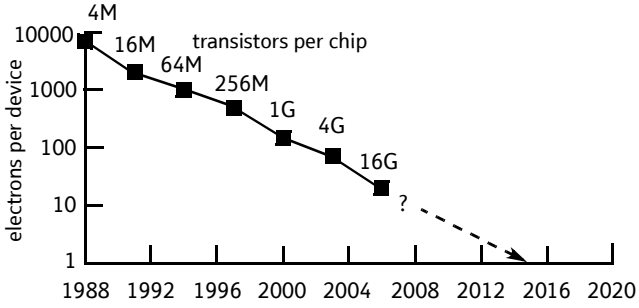

Fig 1 Moore's law (industrial and academic sources) for electrons switched per device, per operation.

There is a semiconductor industry roadmap which clearly indicates a way forward. This is Si-based, but depends both on further implementation of nanotechnology (current transistors already have gate lengths below 100 nm) as well as multicore processor architectures for continued performance gains. Of course the multicore approach of increasing the number of parallel data flows on a single chip does not specifically rely on increasing device speed or reduced device size for better performance. This evolution presents a clear agenda for nanotechnology, starting with the continued refinement of the Si MOSFET (metal-oxide semiconductor field effect transistor), and proceeding through exotic new devices based, for example, on various forms of carbon such as nanotubes or graphene, to schemes which escape the two-dimensionality of current paradigms. There are even more exotic concepts relying on knowledge gained in chemistry and molecular biology to produce self-assembled circuits. Finally, free-space optical and microwave interconnects could liberate circuit designers from the need for continuous physical contact between devices.

However, our view is that multicore architectures are a harbinger of much greater change in IT rather than new devices that are essentially one-for-one replacements of existing devices. At this point chips containing two processing cores are becoming widely available, and very soon more than two will also be commonplace. Designing the necessary parallelism to fully exploit this feature is already very challenging to systems and software designers. Whether by accident or design, the opportunities and challenges presented by multicore, when replicated on the nanoscale, will be much larger and will call for a profound rethink of how we compute and manage data flow. Therefore, we expect the associated nanosystems engineering to have the same impact on systems and software that nanotechnology already has on materials and devices.

The remainder of the present paper provides further detail on new nanodevices, highlighting some that are being considered at the London Centre for Nanotechnology (LCN), which may avoid the physical limitations of current technology. It then points out how the greatest IT needs are actually systems needs, and provides background on how nanotechnology may have an even more profound impact at the systems level than at the device level, where so far it seems to be merely replacing individual devices with presently defined functionalities.

2. Nanotechnology and devices

We are already enjoying the potential of devices with engineered nanoscale features. The best known examples are giant magnetoresistive (GMR) read heads for disk drives. These are metallic multi-layers consisting of superposed thin films of antiferromagnetic, paramagnetic, and ferromagnetic metals. The electrical resistance of these sandwiches varies substantially with magnetic fields, produced for example by the bits encoded in the magnetic domains of a recording disk. Other more futuristic devices, still in research laboratories, include field-effect transistors based on carbon nanotubes, as well as diodes based on other types of nanotubes.

The trend towards ever more exotic building blocks for relatively conventional device ideas is exemplified by Fig 2, showing a four-probe electrical device fabricated by LCN scientists using an amyloid fibril as the fundamental conducting element. Similar fibrils, consisting of misfolded proteins known as prions are implicated in the well-known neurodegenerative Alzheimer's and BSE diseases, and it is possible that their self-assembling tendencies, while potentially fatal in a biological context, might be exploitable in manufacturing processes. In particular, if the separation between contacts is small enough, the fibrils themselves might function as devices with interesting and potentially tuneable electrical conduction. In cases where the fibrils are strongly insulating, their dielectric properties might be of interest for capacitors. They could also serve as templates for more common metals and semiconductors to create MOSFETs in new processes, avoiding some of the lithographic steps (and associated high-precision alignment) currently required.

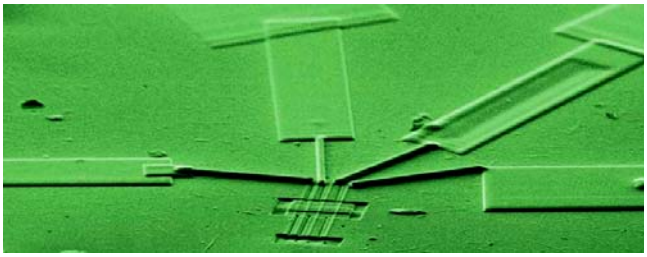

Fig 2 Four electrodes fanning out from a single isolate amyloid fibril [1].

Other LCN research is focused on exploring the fundamental limits of data storage, by attempting to examine the stability of atomic-scale charge and atom configurations. The idea is to neither encode information in magnetic domains nor store charged in some micro- or meso-scopic device, but rather in the charge state of mixed valence ions. Such ions, as for example manganese, can exist stably in several different charge states (e.g. Mn^{2+}, Mn^{3+} and Mn^{4+}), and can be considered to be naturally occurring charge-coupled devices. The technique used in these investigations is scanning tunnelling microscopy (STM), and Fig 3 shows

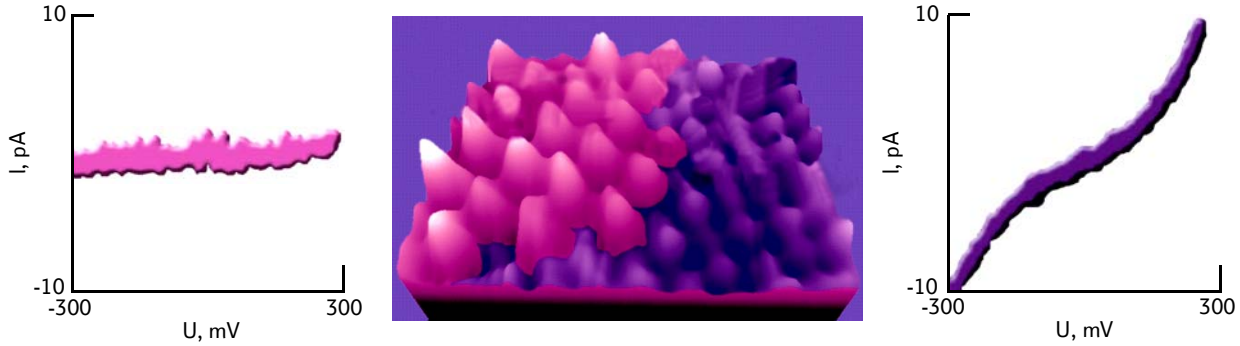

Fig 3 A stable, atomic-scale wall between regions with and without charge order (adapted from Renner et al [2]).

a stable, atomic-scale wall between regions with and without charge order. The material is the perovskite oxide of manganese, $Bi_{1-x}Ca_xMnO_3$. It is apparent that on one side of the wall, the unit cell is the doubled cell of a charge ordered insulator, and on the other, we are left on average with the simple cell associated with metallic behaviour. The image was recorded at room temperature, and so clearly illustrates the potential of storing information on subnanometre scales in solids. Here one might think of bits corresponding to whether there is charge order or not, and this can be read out using either a structural probe, such as STM, or via the tunnelling current which will be higher for small voltages in the metallic, charge disordered regions.

The same experiment also showed that in the metallic phase, the oxygen octahedra surrounding the Mn ions actually tilt in random directions, and that these random tilts are also quite stable at room temperature, which suggests a structural scheme for storing bits on the low Angstrom scale.

It is important to note that while the science underlying single atom data storage at room temperature seems sound, we are not looking at a real technology yet. Realistic writing and read-out schemes are yet to be developed, although the effects that we have imaged may well underlie hysteretic two-terminal oxide device characteristics. These have been studied by others (notably the IBM group of Bednorz [3]) for the order of a decade. In this case, writing is accomplished by the passage of a large current, and read-out simply involves measuring the low current response.

Beyond the search for new devices occupying small volumes, there is also great interest in improving manufacturability by taking advantage of 'natural' chemical trends towards self-assembly. The self-assembly of organic matter is much celebrated and underlies biology, and we have already discussed this in the context of the amyloid fibril of Fig 2. Inorganic matter can also form complex functional aggregates in relatively simple synthetic routes, such as the repeated grinding and firing need to produce ceramics. Figure 4 shows the workings of tunnelling magnetoresistance in a ceramic in the form of a layered oxide of manganese, which in the absence of an applied field is a stack of two-dimensional ferromagnets, arranged such that their magnetic polarisations alternate between layers. The antiferromagnet stacking arrangement blocks the transfer of electrons between the layers, and hence yields a low electrical conductance. Switching to a high conductance state occurs in modest magnetic fields, when the electron gases in adjacent layers acquire parallel ferromagnetic polarisations.

Another important drive in micro- and nano-electronics is towards three-dimensional integration, with electrons confined to planar racetracks with well-specified shunts between them. In the first instance the shunts need only be passive, since active devices controlling access to them can be incorporated within the planes. If the planes can be within nanometres of each other, a three-dimensional integrated circuit with a thickness of a millimetre would have the power of a million (1 mm/1 nm) circuits today, provided of course that many practical issues, mainly surrounding the correspondingly thin gate oxides, could be dealt with! One good starting point, however, would be a material where the electrons (or holes) are naturally confined to layers, with low dissipation for flow within the layers, but a large barrier to overcome to switch between layers, unless a shunt is encountered.

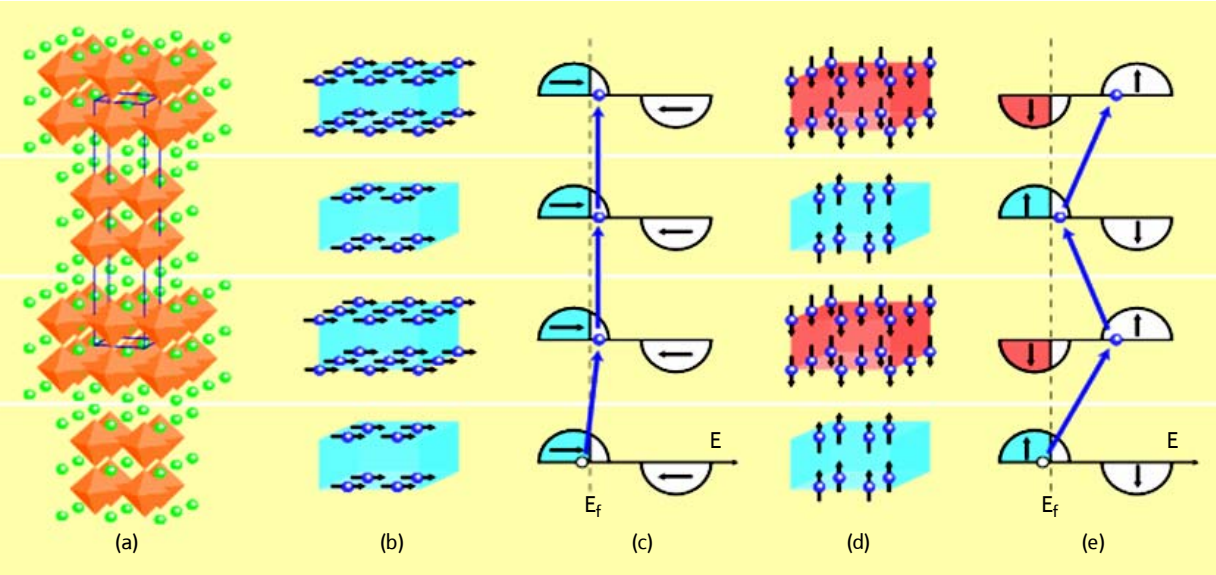

Fig 4 Self-assembled stack of tunnelling spin valves ($La_{2-2x}Sr_{1+2x}Mn_2O_7$ with $x = 0.3$) (from Perring et al [4]).

It turns out, not surprisingly, that the same layered material, which exhibits self-assembled spin valves, also displays confinement. The reason for this is that the charge carriers in the planes are surrounded by texture in the lattice and magnetic degrees of freedom. These textures, referred to as 'polarons', are due to the distorting influence on their surroundings of the electrons or holes at the polaron 'nuclei'. As desired they slide easily in the planes, but need to be broken up and reconstructed if they are to travel between planes. Occasionally, they are trapped at defects, and can then be observed as static objects in scanning tunnelling micrographs, as in Fig 5.

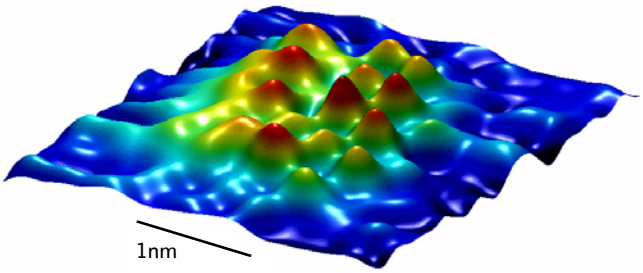

Fig 5 Trapped 'polaron', the texture surrounding the charge carriers in the planes of $La_{2-2x}Sr_{1+2x}Mn_2O_7$ with $x = 0.3$, a ceramic layered material exhibiting two-dimensional confinement of charge carriers and therefore a prototype host for three-dimensional integrated electronics (adapted from Ronnow et al [5]).

The same structural archetype as in Fig 4 also has a superconducting realisation, where copper oxide bilayers are weakly (Josephson) coupled along the stacking direction. Using focused ion beam technology (see Fig 6), it is possible to isolate of the order of 100 of these naturally self-assembled superconducting electronic devices. We have carried out experiments to observe the associated latching behaviour that makes them interesting for potential future applications. The multilevel latching displayed by the data is derived from the mesoscopic number of junctions. Beyond yielding an interesting new device type, the experiments represent one of the most direct confirmations that high-temperature superconductors can be thought of as Josephson-coupled two-dimensional superconductors.

3. Nanotechnology and systems

In section 1, we described the physical limitations which will soon be hampering further 'Moore's Law' growth if we simply continue to refine existing (largely MOSFET-based) technology. In parallel, there are also ever more challenging systems limitations. Succeeding generations of processors and software are growing in cost and complexity. One obvious result is more late deliveries and more unreliable products.

Just as the fundamental enabler of hardware growth has been the transistor, the underlying paradigm for software is the Turing/von Neumann machine, not coincidentally discovered at around the same time at the end of World War II. In a Turing/von Neumann machine, stored instructions are executed sequentially on stored data.

The research described in section 2 will naturally yield devices that would slot into Turing/von Neumann machines. Arguably more significant developments would lead not only to new devices, but would deal with the systems challenges of the Turing/von Neumann paradigm. LCN researchers are active in this area as well, most notably in efforts to bring quantum information processing into reality. Here, interacting quantum bits naturally allow the parallelisation needed to solve hard problems, such as prime factorisation. Also, quantum tunnelling permits new search strategies, which are self-regulating in a meaningful sense and are therefore easily programmed. Hardware implementations would permit the elimination of complex software programs, which have to be translated into machine instructions via nested sequences of equally complex compilers, as well as yielding dramatic

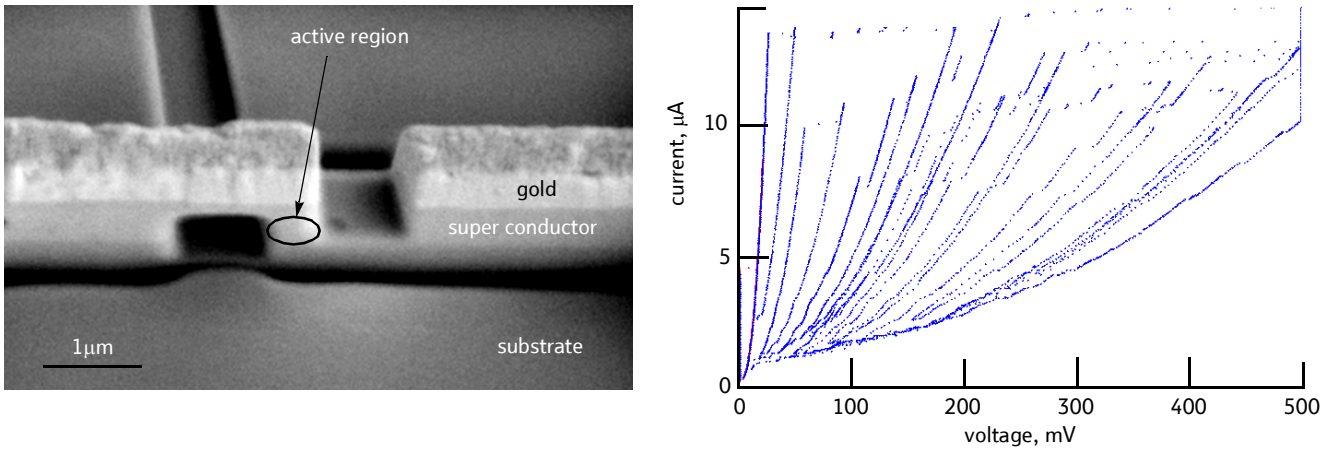

Fig 6 Self-assembled stacks of Josephson junctions in $Tl_2Ba_2CaCu_2O_8$ [6].

improvements in performance for certain problems. Such hardware implementations appear feasible given demonstration experiments [7,8] that LCN scientists are involved with, as well as theoretical and numerical work done by others which has been motivated by these experiments.

A particularly fruitful concept has been the idea that a computation or decision that involves many degrees of freedom can be translated into an optimisation problem, where a function of many variables is minimised, subject to certain constraints. The minimum can then be sought using a stochastic sampling process, analogous to that followed by thermodynamic systems being cooled under near-equilibrium conditions.

The technique is call 'simulated annealing', because temperature is being added as an artificial parameter to allow the exploration of the space of possible solutions to the problem. This space of possible solutions is called the 'phase space'. Instead of developing a conventional algorithm for searching the phase space, 'temperature' is introduced to weight state probabilities. The system is 'warmed up' and 'cooled' slowly to find the minimum, exactly as happens for natural systems settling into their ground states. This approach was first taken over two decades ago for finding good (but not necessarily the best) solutions to a variety of scheduling and resource allocation problems, many related to that of the famous travelling salesman who needs to maximise exposure to clients in a network of cities while minimising the cost and time of the journey, with competing constraints.

More recently, quantum tunnelling (see Fig 7) rather than thermal barrier hopping has been added as a parameter for moving through phase space. In certain instances it has been shown to produce different, and probably better, results than simple thermal annealing, even when implemented on a conventional computer. The promise of nanotechnology is that it could permit

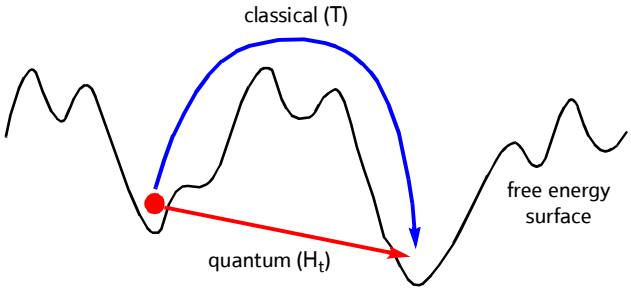

Fig 7 Schematic showing thermal and quantum annealing protocols for finding the optimum (lowest value) of an objective function, as defined — for example — for a communications routing problem (from Brooke et al [7]).

direct physical implementation of device networks; it would be configured to find optima via 'natural' (quantum and thermal) means rather than either conventional sequential computation, or simulation — with the help of a random number generator — on a conventional computer. A schematic of what we have in mind is shown in Fig 8, the left-hand side of which shows the underlying (Feynmann) view of a computation, and the right hand side its implementation in a real physical system. In the real physical implementation, the starting point is a particular state described by an initial wavefunction, and output is represented by a physical measurement $\chi(\omega)$ on the final wavefunction, which is reached via a set of quantum and thermal annealing steps.

A related concept is parallelism, and its associated difficulties and opportunities. As devices become smaller and more tightly packed with the advances in nanotechnology, interactions between them become increasingly unavoidable. We will then be forced into accepting that operations on a circuit chip will be difficult to view as independent operations. The most extreme form of such behaviour is quantum entanglement. This means that physically separated bits of information are correlated even when there is no classical interaction between them, and leads to what we might consider 'inevitable parallelism'.

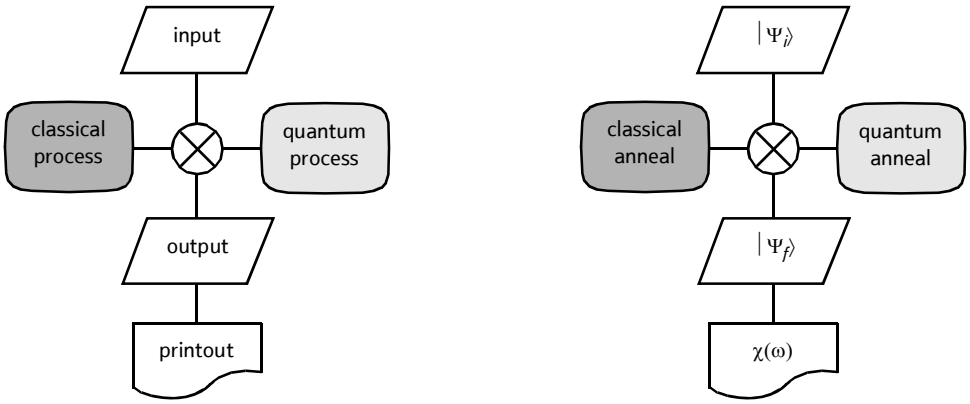

Fig 8 Schematic of general quantum/classical computer as viewed, at left, by a computer scientist and, at right, by a physicist or hardware engineer (see text for more detail).

Quantum computing aims to harness entanglement. The resolution of certain hard problems, such as the prime factorisation of large numbers using quantum gates, has been the focus of huge activity over the last decade or so. The potential that quantum computers might actually break current security codes has provided a major incentive for such work. The key technical issue has been the controlled entanglement of quantum bits, or 'qubits', and thus far, the record is held by an Austrian ion trap, which contains seven 'qubits'.

Quantum computing is to be distinguished from quantum communications, which is much closer to market (indeed, there are products already available for secure communications). This relies on the principle that observation, i.e. eavesdropping, alters quantum states, and therefore marks a message as having been compromised.

For many obvious reasons, the search for realistic solid state implementations of quantum computing continues, and there are programmes to harness elements as disparate as superconducting rings, single impurities in semiconductors, and one-electron transistors as hosts for qubits. The tools of nanotechnology are essential for such implementations as they provide the underlying metrology and fabrication capabilities. It would obviously be highly interesting to discover a room-temperature scheme, reliant only on relatively minor modifications of current silicon integrated circuit technology. The authors of this paper are involved with one candidate proposal of this variety, illustrated in Fig 9.

The key idea is to use the electron spins of randomly dispersed impurities in silicon as qubits, and to use the magnetic interactions between them to perform entangling operations. The interactions are controlled externally using optical excitation of 'control bits', namely the atomic orbitals of other impurity atoms situated between the 'qubits'. The distances between entangled impurities in this scheme, as well as the distances between the qubits and control bits, are all in the nanometre to several nanometre range. However, they do not need to be dispersed deterministically as long as appropriate nanometrology is available to discover where they are actually located, so as to specify laser addressing protocols (including wavelengths, and pulse shaping and timing) for both reading and writing.

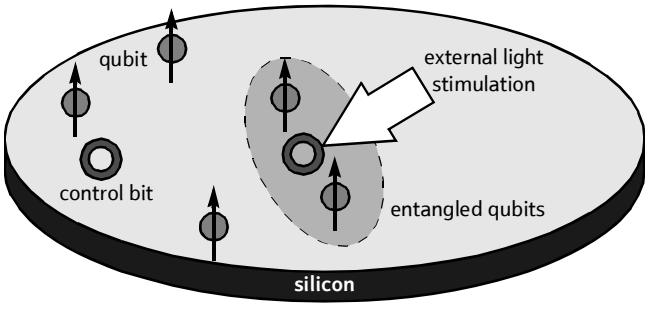

Fig 9 Schematic of a possible silicon-based implementation of a quantum computer. Spin 'qubits' are associated with the impurity atoms, and their states and quantum entanglement are controlled via optical excitation of 'control bits' associated with the electronic orbitals of the darker impurity atoms (adapted from Stoneham et al [9]).

4. Conclusions

Nanotechnology underpins the current push to reduce the dimensions of conventional electronic devices. At the same time, nanoscience is providing new concepts for devices which could replace the current components of today's electronic circuits. Further into the future, we see nanotechnology as defining not just new devices, but also new computation and communications paradigms, of which the most prominent so far are quantum computing and cryptography.

What we are holding out for is replacing current hardware with new networks of devices with collective intelligence that will solve not only the seemingly mundane problems of energy consumption and fault tolerance, but also the larger problem of reliable systems programming.

Indeed, just as the developments leading to the invention of the transistor coincided with the formulation of the Turing/von Neumann paradigm for serial stored program machines, novel nanotechnology-enabled hardware should inspire and coexist with new concepts for management of computation and communications. In this broader context the impact of nano-systems science is likely to eclipse that of nanotechnology-based hardware itself.

References

1 Warburton P, McKendry R, Mesquida R and Aeppli G: Unpublished work.

2 Renner C et al: 'Atomic-scale images of charge ordering in a mixed-valence manganite', Nature, 416, pp 518—521 (2002).

3 IBM Research — http://www.research.ibm.com/

4 Perring T G et al: 'Ordered stack of spin valves in a layered magnetoresistive perovskite', Physical Review B, 58, No 22 (1998).

5 Ronnow H M et al: 'Polarons and confinement of electronic motion to two dimensions in a layered manganite', Nature, 400, pp 1025—1028 (April 2006).

6 Warburton P: Unpublished work.

7 Brooke J et al: 'Quantum annealing of a disordered magnet', Science, 284, pp 779—781 (April 1999).

8 Ghosh S et al: 'Entangled quantum state of magnetic dipoles', Nature, 425, pp 48—51 (2003).

9 Stoneham A M et al: 'Optically driven silicon-based quantum gates with potential for high-temperature operation', J Phys Condens Matter, 15, No 27, pp L447—L451 (July 2003).

Gabriel Aeppli is the Quain Professor of Physics and Head of the Condensed Matter Physics Group at UCL, and the Director of the London Centre for Nanotechnology.

Prior to taking up these posts in the autumn of 2002, he was a Senior Research Scientist for NEC (Princeton), a Distinguished Member of Technical Staff at Bell Laboratories, a Research Assistant at MIT, and an industrial co-op student at IBM.

He obtained a BSc in Mathematics and PhD, MSc and BSc in Electrical Engineering from MIT. Honours include the Oliver Buckley Prize of the American Physical Society (2005), the Neel Medal and Magnetism Prize of the International Union of Pure and Applied Physicists (2003), Royal Society Wolfson Research Merit Award (2002), Mildner Lecturer, Department of Electronic and Electrical Engineering UCL (2002), Fellow of the American Physical Society (1997), Fellow of the Japan Society for the Promotion of Science (1996).

In addition, he has been a member and chairman of many panels, sponsored by the USDOE, American Physical Society, EPSRC, and National Research Council (USA), among others. His main research interests are quantum information processing and nanotechnology, including especially its manifestations in the life sciences.

Paul Warburton is the group leader of the UCL Quantum Nanoelectronics Group and a Senior Lecturer in Electrical Engineering. He first established the group in 1995 at King's College London.

Following notable research successes in both low-loss tuneable microwave devices and Josephson junction technology, he and the group relocated to University College London in August 2001, and, in 2006, to the newly built London Centre for Nanotechnology.

Christopher Renner is a Principal Research Fellow in the London Centre for Nanotechnology and the Department of Physics and Astronomy, University College London. He obtained his PhD in Physics in 1993 at the University of Geneva in Switzerland. Before joining UCL in 2002, he spent four years as a Visiting Scientist at NEC in Princeton (USA). He has much experience designing and using ultra-high vacuum variable temperature STMs in high magnetic fields, mostly in the study of strongly correlated electron materials. He has also been a Business Fellow for the London Technology Network since 2003, developing interactions between the University and the Industry. He received the 1994 Jean Wurth Award from the University of Geneva and the 1998 IBM award from the Swiss Physical Society for his pioneering STM studies of high-temperature superconductors.