A Device Architecture for Computing with Quantum Dots

CRAIG S. LENT AND P. DOUGLAS TOUGAW

Invited Paper

We describe a paradigm for computing with interacting quantum dots, quantum-dot cellular automata (QCA). We show how arrays of quantum-dot cells could be used to perform useful computations. A new adiabatic switching paradigm is developed which permits clocked control, eliminates metastability problems, and enables a pipelined architecture.

Keywords— *Molecular electronics, quantum-effect semiconductor, devices, quantum-well devices, single-electron devices.*

I. INTRODUCTION

Recent years have seen a dramatic improvement in the size and speed of electronic devices; the exponential pace of microelectronics is well known. Although current trends may continue for some time, inevitable road blocks loom. Whether or not one can predict with confidence how long the exponential path can be extended, it makes sense to now explore more radical technologies that could leapfrog conventional CMOS and enable scaling to continue unhindered down to molecular sizes.

It is helpful to appreciate that current MOS transistors are direct descendants of the electromechanical switches first used by Zuse to code digital information in an electronic form. Representing binary information by turning on or off a current switch has been one of the most fruitful ideas in the history of technology. This paradigm does however have serious drawbacks as device sizes are reduced. The interconnect problem is one. One needs to distribute signals over large distances, which involves charging long lines. Remarkable complexity attends the routing of signals on multiple levels. At the other end, as transistors become smaller, the quantization of charge both in the channel and in the doping layer become significant. When reduced to nanometer scales, current switches may not be the best way to code information.

The developing technology of quantum dot fabrication may prove a key element in crafting another approach. Researchers have already demonstrated the ability to fabricate quantum dots with single charges [1], to make large arrays of dots and control their occupancy [2],[1] and to put dots in close enough proximity to observe both tunneling and Coulomb coupling between dots [3]–[6].[2] Of course, these experiments are difficult and consistent, high-yield processing techniques are not yet at hand.

Here we will assume that the fabrication problems can be solved, and explore the possibilities this quantum dot nanofabrication technology might open were it to succeed. This in no way minimizes the difficulties associated with the fabrication process itself. It may be that some of those difficulties will actually prove insurmountable in the long run, or it may be that simple reliable fabrication techniques are close at hand. In either case, addressing how quantum dots might be used in concrete ways to accomplish computing will serve both to motivate the search for solutions to fabrication problems and to focus on those fabrication requirements that are most important.

We discuss an approach to computing with quantum dots, the quantum-dot cellular automata (QCA), which is based on encoding binary information in the charge configuration of quantum-dot cells [7]–[17]. The interaction between cells is Coulombic, and provides the necessary computing power. No current flows between cells and no power or information is delivered to individual internal cells. Local interconnections between cells are provided by the physics of cell–cell interaction.

The next section describes the QCA cell and the process of building up useful computational elements from it. The discussion is mostly qualitative and based on the

Manuscript received August 2, 1996; revised January 21, 1997. This work was supported in part by the Advanced Research Projects Agency and the Office of Naval Research, and based in part upon work supported under a Center for Applied Mathematics Graduate Fellowship.

C. S. Lent is with the Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556 USA (e-mail: lent@nd.edu).

P. D. Tougaw is with the Electrical and Computer Engineering Department, Valparaiso University, Valparaiso, IN 46383 USA (e-mail: dtougaw@diamond.valpo.edu).

Publisher Item Identifier S 0018-9219(97)02731-X.

[1] The ability to control quantum dot occupancies over as many as 10^8 dots using a back-gating technique has been reported by Meurer *et al.* [2].

[2] For a recent collection of quantum dot papers, see [6].

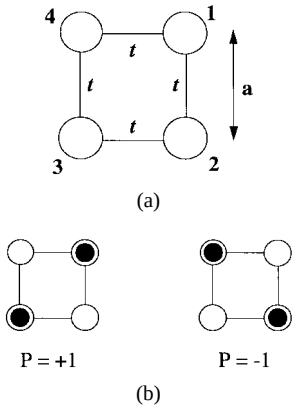


Fig. 1. Schematic of the basic four-site cell. (a) The geometry of the cell. The tunneling energy between two neighboring sites is designated by t, while a is the near-neighbor distance. (b) Coulombic repulsion causes the electrons to occupy antipodal sites within the cell. These two bistable states result in cell polarizations of $P = +1$ and $P = -1$ [see (6)] [11].

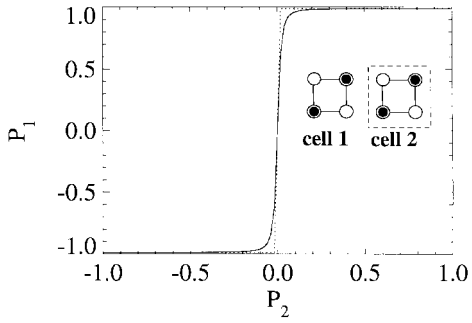


Fig. 2. The cell–cell response. The polarization of cell 2 is fixed and its Coulombic effect on the polarization of cell 1 is measured. The nonlinearity and bistable saturation of this response serves the same role as gain in a conventional digital circuit [11].

intuitively clear behavior of electrons in the cell. Section III discusses the quantum mechanical description of the cell and describes the way in which more detailed and rigorous calculations of QCA array behavior can be obtained. Section IV, the main thrust of this paper, is a discussion of how switching of cellular arrays can be accomplished. We focus particularly on a newly developed paradigm of adiabatically clocked arrays. Section V extends the implications of this approach to a pipelined architecture. In Section VI we mention various alternatives for QCA fabrication and end with some conclusions.

II. QCA Overview

We begin with a qualitative overview of how QCA cells work and interact to form a computational architecture. The next section will discuss the quantum mechanical description appropriate for a full exploration of the physics of the cellular arrays. Here we will rely on a more intuitive description.

A schematic diagram of a four-dot QCA cell is shown in Fig. 1(a). The cell consists of four quantum dots positioned at the corners of a square. The cell contains two extra mobile electrons, which are allowed to tunnel between neighboring sites of the cell. The compensating positive charge is fixed and immobile [12]. Tunneling out of the cell is assumed to be completely suppressed by the potential barriers between cells. It is also possible to add a fifth dot at the center of the square; the addition of this dot improves the behavior of the cell slightly but for simplicity we will focus mainly on the four-dot cell.

If the barriers between cells are sufficiently high, the electrons will be well localized on individual dots. The Coulomb repulsion between the electrons will tend to make them occupy antipodal sites in the square as shown in Fig. 1(b). For an isolated cell there are two energetically equivalent arrangements of the extra electrons which we denote as a cell polarization $P = +1$ and $P = -1$. The term "cell polarization" refers only to this arrangement of charge and does not imply a dipole moment for the cell. The cell polarization is used to encode binary information—$P = +1$ represents a binary 1 and $P = -1$ represents a binary 0.

The two polarization states of the cell will not be energetically equivalent if other cells are nearby. Consider two cells close to one another as shown in the inset of Fig. 2. The figure inset illustrates the case when cell 2 has a polarization of $+1$. It is clear that in that case the ground-state configuration of cell 1 is also a $+1$ polarization. Similarly if cell 2 is in the $P = -1$ state, the ground state of cell 1 will match it. The figure shows the nonlinear response of the cell-cell interaction to which we will return in the next section.

A QCA wire is shown in Fig. 3(a). The left-most cell is fixed with a polarization representing the input. The ground state configuration of the remaining free cells is then one with each cell polarized in the same way as the input cell. We can consider this transmission of the input signal from one end to the other (again, the precise temporal meaning of "transmission" in this case will be discussed below).

Cells which are positioned diagonally from each other tend to anti-align. This feature is employed to construct an inverter as shown in Fig. 3(b). The anti-alignment can also be seen by examination to be a simple consequence of the mutual repulsion between electrons and the geometry of the cells. Although two diagonal cells function as an inverter, this more symmetric design ensures exact symmetry between the inversion of a one and a zero. Fan-out of a signal is illustrated in Fig. 3(c).

Fig. 3(d) shows the fundamental QCA logical device, a three-input majority gate, from which more complex circuits can be built. The central cell, labeled the device cell, has three fixed inputs, labeled A, B, and C. The device cell has its lowest energy state if it assumes the polarization of the majority of the three input cells. The output can be connected to other wires from the output cell. The difference between input and outputs cells in this device, and in QCA arrays in general, is simply that inputs are fixed and outputs are free to change. The inputs to a particular device can come from previous calculations or be directly fed in from array edges. The schematic symbol used to represent such a gate is also shown in Fig. 3(d).

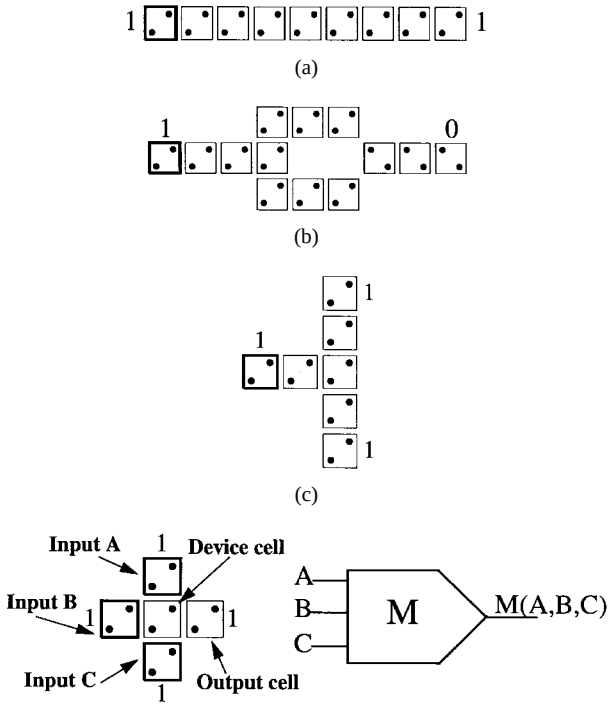


Fig. 3. Fundamental QCA devices: (a) The binary wire allows transmission of information from one point to another within the array, (b) the inverter uses diagonal anti-voting behavior to invert the signal, (c) fanout allows the result of a calculation to be propagated to two or more other points within the array, and (d) the majority logic gate, the fundamental logical element of a QCA array, and its logical symbol [11].

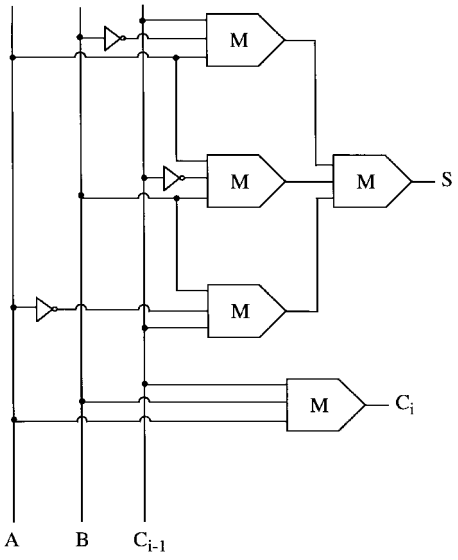


Fig. 4. Schematic diagram of the logic gates needed to make a QCA single-bit full adder. This implementation requires five majority logic gates and three inverters.

It is possible to "reduce" a majority logic gate by fixing one of its three inputs in the 1 or 0 state. If the fixed input is in the 1 state, the OR function is performed on the other two inputs. If it is fixed in the 0 state, the AND function is performed on the other two inputs. In this way, a reduced majority logic gate can also serve as a programmable AND/OR gate. Combined with the inverter shown above, this AND/OR functionality ensures that QCA devices provide logical completeness.

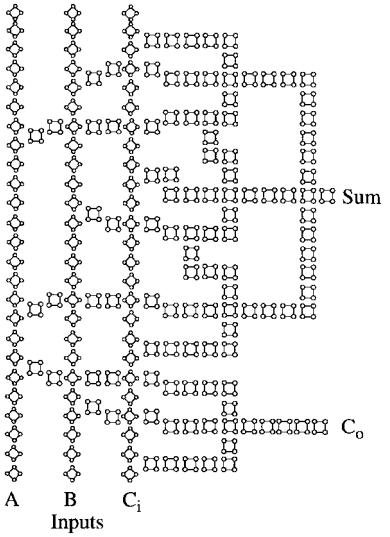


Fig. 5. Schematic diagram of the QCA cell layout necessary to implement the logic gates shown in Fig. 4. The necessary wire crossings are implemented by the three parallel lines of rotated cells, which distribute the inputs to the two levels of majority gate logic.

As an example of more complex QCA arrays we consider the implementation of a single-bit full adder. A schematic of the logic device layout for an adder implemented with only majority gates and inverters (using the conventional symbol) is shown in Fig. 4. Fig. 5 shows the schematic layout of QCA cells necessary to implement the design of Fig. 4. The five majority gates in Fig. 4 are seen to be realized in Fig. 5 using three-input junctions of wires. A full quantum mechanical simulation of such an adder has been performed [11], and verifies that it yields the correct ground state output for all eight possible combinations of the three inputs.

The adder layout in Fig. 5 illustrates another useful feature of QCA design, namely the ability to cross wires in the plane. In conventional technology, information is coded in voltages or currents in conductors and wire crossings require a bridge out of the plane. Using QCA wires, we have been able to show that signals can be crossed in a coplanar way by employing rotated versions of the QCA cells. The vertical lines on the left of Fig. 5 can be seen to be wires of such rotated cells. Being diagonally oriented, their polarizations alternate down the line (an inverter chain). The advantage of using such lines to distribute the signals is that the usual wires, as in Fig. 3(a), can cross the rotated wires without interference or crosstalk. Again, this feature is a consequence simply of the Coulomb interaction and the symmetries present in the cell charge. The alternating polarization in the rotated wires also permits easy extraction of both the signal and its complement.

Without getting too far into implementation-specific features, let us briefly address the question of input and output in a QCA array. Setting an input wire requires coercively

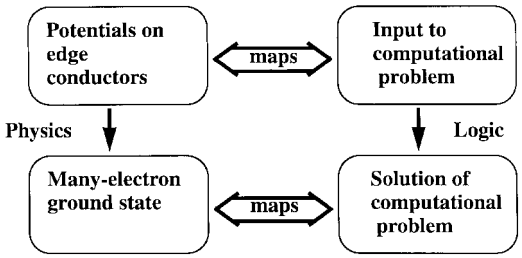


Fig. 6. Schematic of the mapping of device physics onto computational logic. By controlling the geometry of the device, it is possible to use this mapping to perform useful calculations.

setting the state of the first cell in the wire. This can be accomplished very simply by charging nearby conductors to repel electrons from one dot and attract them to another. In quantum dots made in semiconductors, this has become a standard experimental technique, usually called a "plunger electrode," to alter electron occupancy of a dot [3]–[5]. Reading an output state is more difficult. We require the ability to sense the charge state of a dot without having the measurement process alter the charge state. Since the local charge produces a local electrostatic potential, this is really a question of constructing a small electrometer. Fortunately, electrometers made from ballistic point-contacts and from quantum dots themselves have already been demonstrated. These electrometers can noninvasively measure the charge state of a single dot [18]. Note that input and output are only performed at the edges of the array; no information or energy need flow to interior cells.

A QCA array like the adder discussed here works because the layout of the quantum-dot cells has provided a mapping between the physical problem of finding the ground state of the cells and the computational problem. The physical problem can be stated as follows: Given the boundary conditions imposed by the input, what is the lowest energy configuration of the electrons in the cellular array? It is the ability to make this mapping between the physical ground state and the unique logical solution state that is at the heart of the QCA approach.[3] This is illustrated schematically in Fig. 6.

Before turning to the details of the quantum mechanical description of QCA arrays, we can now summarize by enumerating the general features of the QCA paradigm:

- bistable cells code bit information in their internal configuration;
- physically mediated cell-cell interaction provides coupling between the states of nearby cells;
- inputs to array are set by physically coercing edge cells to particular states;
- outputs are read by noninvasively sensing the state of edge cells;
- computing is accomplished by the mapping between the physical ground state of the array and the logical solution state of the computational problem.

[3]In almost all cases explored, the obvious logical layout procedure results in a correct mapping between ground state and solution state. Some pathological cases have been constructed, so in general a verification of this is necessary.

III. Modeling Quantum Cellular Automata

A. A Quantum Description of the QCA Cell

A simple Hamiltonian of the extended-Hubbard type is used to describe the QCA cell shown in Fig. 1. We ignore internal degrees of freedom of the cell, treating each quantum dot as a site. The Hamiltonian used to model the cell is (see (1) at the bottom of the page). Here we use the usual second-quantized notation where $\hat{a}_{i,\sigma}$ ($\hat{a}^{\dagger}_{i,\sigma}$) annihilates (creates) an electron on site i with spin σ. The number operator for electrons of spin σ on site i is $\hat{n}_{i,\sigma} = \hat{a}^{\dagger}_{i,\sigma}\hat{a}_{i,\sigma}$. In (1), the first term represents the on-site energy of each dot. The potential energy of an electron at dot i due to charges outside the cell (including charges in other cells) is V_i. The second term accounts for electron tunneling between sites, with $t_{i,j} = t$ for neighboring sites and $t_{i,j} = 0$ for antipodal sites. The third term is the on-site charging cost to put two electrons of opposite spin on the same dot, and the last term corresponds to the Coulombic interaction between the electrons on different sites within a cell.

To find the stationary states of the cell, we solve the time-independent Schrödinger equation

$$\hat{H}^{cell}|\psi_i\rangle = E_i|\psi_i\rangle \tag{2}$$

where $|\psi_i\rangle$ is the ith eigenstate of the Hamiltonian, and E_i is the corresponding eigenvalue. These eigenstates are found using the many-particle site-ket basis for four sites and two electrons of opposite spins:

$$\begin{aligned} |\phi_1\rangle &= \left|\begin{matrix}0&0&0&1\\0&0&0&1\end{matrix}\right\rangle, \\ |\phi_2\rangle &= \left|\begin{matrix}0&0&0&1\\0&0&1&0\end{matrix}\right\rangle, \cdots, \\ |\phi_{16}\rangle &= \left|\begin{matrix}1&0&0&0\\1&0&0&0\end{matrix}\right\rangle. \end{aligned} \tag{3}$$

In this notation the columns correspond to site (dot) indices and the rows correspond to spin (upper row for spin up). We calculate the Hamiltonian matrix in this basis set by numerically evaluating each matrix element

$$H_{ij} = \langle\phi_i|\hat{H}|\phi_j\rangle \tag{4}$$

$$H^{cell} = \sum_{i,\sigma}(E_0 + V_i)\hat{n}_{i,\sigma} + \sum_{i>j,\sigma} t_{i,j}(\hat{a}^{\dagger}_{i,\sigma}\hat{a}_{j,\sigma} + \hat{a}^{\dagger}_{j,\sigma}\hat{a}_{i,\sigma}) + \sum_{i} E_Q\hat{n}_{i,\uparrow}\hat{n}_{i,\downarrow} + \sum_{i>j,\sigma,\sigma'} V_Q\frac{\hat{n}_{i,\sigma}\hat{n}_{j,\sigma'}}{|R_i - R_j|} \tag{1}$$

and finding the eigenvectors of the resulting 16×16 matrix.

The ground state of the cell, $|\psi_0\rangle$, is represented in this basis as

$$|\psi_0\rangle = \sum_j \psi_j^0 |\phi_j\rangle. \tag{5}$$

Here, $|\phi_j\rangle$ is the jth basis vector and ψ_j^0 is the coefficient of that basis vector, which is found by direct diagonalization of the Hamiltonian.

If the tunneling between sites of the cell is relatively weak (high tunnel barriers), the electron number on each site will be approximately quantized. In this case, it is qualitatively clear that the ground state of such a cell would resemble one of those shown schematically in Fig. 1(b). If the tunneling between sites of the cell is increased, the localization of electrons on each site will be reduced. If the tunneling energies become comparable to the Coulombic energies associated with the cell, the localization of the two-electron wavefunction is removed, and the polarization of the cell is eliminated. As long as the tunneling energies within the cell are significantly less than the Coulombic energies associated with the cell, the electrons will remain largely localized, and cells will be well polarized. This is the case for the cell proposed above.

In order make quantitative this notion of polarization we define the cell polarization, which is a property of the ground state eigenfunction $|\psi_0\rangle$, as follows:

$$P \equiv \frac{(\rho_1 + \rho_3) - (\rho_2 + \rho_4)}{\rho_1 + \rho_2 + \rho_3 + \rho_4}. \tag{6}$$

Here, ρ_i is the expectation value of the number operator on site i for the ground state eigenfunction

$$\rho_i = \langle \psi_0 | \hat{n}_i | \psi_0 \rangle. \tag{7}$$

B. Quantifying the Cell–Cell Interaction

We can now consider the interaction between two cells in a quantitative way. We consider two cells whose centers are separated by $3a$ (a is the near-neighbor distance between dots). We calculate the polarization induced on one of the cells by a polarization of the other. As seen in the schematic inset to Fig. 2, we fix the polarization of cell 2 at a series of values ranging from -1 to $+1$ and determine the Coulombic effect this fixed charge density arrangement has on cell 1. Including the appropriate induced potential in the on-site energy of each dot in cell 1, we then solve the Schrödinger (2) directly and determine the quantum mechanical ground state of cell 1. From this state the polarization of cell 1 can be calculated using (6) and (7). The results of this calculation are shown in Fig. 2, which we refer to as a cell–cell response function.

The important feature of Fig. 2 that the cell–cell response of a typical QCA cell is highly nonlinear and bistable. Since even a small polarization of cell 2 leads to an almost complete polarization of cell 1, this bistable saturation behavior provides the analogue of *gain* in conventional digital logic devices. Degradation in the polarization of a cell due to fabrication-related imperfections are rapidly corrected in subsequent cells of the array. The nonlinear response means that signal levels are restored at each stage of the calculation. (See below Section III-C.)

C. Examples of Modeling QCA Devices

We have solved the Schrödinger equation for several QCA arrays to verify that the physical ground state and the computational solution state do indeed match. For arrays with a few cells a straightforward extension of the technique described in Section II-A above can be used. The basis set for each cell in the array is constructed and a direct product space is constructed from these to describe the array as a whole. If each cell requires a basis set of 16 kets, then the direct product space for N cells has dimension 16^N. We have used two approaches to enable calculations for larger arrays. One approach is to construct an optimized two-dimensional (2-D) basis set for each cell [15], thus reducing the direct product space to dimension 2^N. For larger arrays this is still intractable and we have adopted a technique which we term the intercellular Hartree approximation [10]. In this approach the ground state is calculated iteratively by solving each cell exactly, using (2), and updating the effects of that cell on the potential energies in all other cells. This approach matches very well with the more exact direct product diagonalization when applied to smaller arrays.

For concreteness, we have chosen a model "standard semiconductor cell" with parameters appropriate to what we estimate to be the limits of semiconductor quantum dot fabrication using electron beam lithography. We choose a cell with a near-neighbor dot distance $a = 20$ nm. Cells are spaced with centers separated by $3a$. The tunneling energy between neighboring dots within the cell will be $t = 0.3$ meV, and all other physical constants will match those of GaAs. The tunneling energy t represents the effect of the potential barriers between the dots. A full exploration of the effect of varying this parameter shows that as long as t is small enough (i.e., the potential barriers are high enough) to make the Coulombic terms in the Hamiltonian dominate over the kinetic energy (t-related) terms, the precise value of t is not critical. (Varying the potentials through this range will be important in adiabatic switching discussed in Section IV.)

Fig. 3 illustrates the results of these ground-state calculations for the devices we have discussed—a wire, inverters, fan-out, and majority gate. Although in the discussion in Section II we treated these figures as simply schematic, they are actually more than that. The diameter of each dot shown is proportional to the calculated charge on each site. (These figures actually reflect the result for cells with a fifth central dot.) The cells with darker borders are the driver cells (inputs), and have a fixed polarization that is included in the self-consistent calculation of the states of the other cells. Cells with lighter borders are standard QCA cells and are free to react to the polarization of the driver cells as well as the polarization of their neighbors. Similar calculated results for the wire crossing and adder can be found in [11].

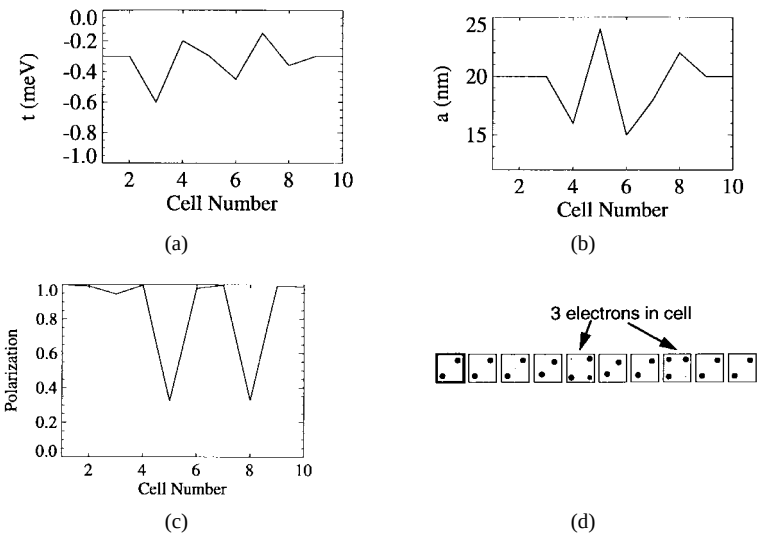


Fig. 7. Robustness of the wire response. The bistable saturation of the cell–cell response leads to very robust behavior of the binary wire: (a) randomly varying tunneling energy of cells in a binary wire, (b) randomly varying size of cells in a wire, and (c) the resulting polarization of the binary wire. The large drops in polarization of cells 5 and 8 are due to the presence of an additional electron in those cells, as shown in (d). Local decreases in polarization are rapidly recovered by the bistable saturation of the next cell in the line.

Fig. 7 shows the calculated results for a disordered wire. The tunneling energies and dot sizes are varied randomly down the wire. In two cases the severe error of having an extra electron in the cell occurs. Nevertheless the wire functions properly, faithfully transmitting either input state. The highly nonlinear response function acts to constantly correct mistakes and restore the signal level. Of course, there are limits to this self-correction and sufficiently severe variations or damage will pin particular cells in one state and destroy the device behavior.

D. Thermodynamic Considerations

QCA arrays operate, as we have seen, by a mapping between the ground state of a physical array and the solution state of a computational problem. Using the ground state to accomplish the computing is really a concession to the difficulty of making separate connections to each cell. With separate power connections, power could flow to keep the system away from equilibrium. However, the cost in interconnection complexity as well as power dissipation is just too high. After all, the idea is that we should be able scale this architecture down to molecular lengths.

Computing with the ground state has the attendant difficulty of being temperature sensitive. If thermal fluctuations excite the array above its ground state, wrong answers can appear at the outputs. To be robust, the excitation energy must be well above k_BT. We can quantify the effect of temperature by replacing the expectation value in (7) with a thermal expectation value, including both the quantum effects and the trace over the full density matrix to include the effect of all thermally excited states [14]. The results give a maximum operating temperature for cells which depends on the size of the cell. As cell size decreases, the energy separations between states increase and higher temperature operation is possible. For the *standard semiconductor cell* (described in Section II-C above) we find that cells work up to about 7 K; at higher temperatures the cell–cell response function becomes nearly linear. To point the direction for future scaling, we consider a *macromolecular cell*, whose near-neighbor distance is reduced to 2 nm with a relative dielectric constant of unity. In this case the maximum operating temperature increases to 700 K.

Thermodynamic limitations impose themselves in another way to limit the practical size of individual arrays. Consider for simplicity a linear array of N cells acting as a wire transmitting a logical one. The ground state consists of all the cells with the same polarization as the input. The first excited state of the array consists of the first m cells polarized in the 1 state and $N-m$ cells in the 0 state. Let us call the excitation energy of this state E_k, the energy of introducing a "kink" in the polarization. The energy is independent of where the kink occurs, i.e., the precise value of m. As the system size N becomes larger, the kink energy E_k remains the same, however the entropy of this excited state increases—there are more ways to make a mistake in a larger array. When the array reaches a certain size, the free energy of the mistake state becomes lower than the free energy of the correct state. A complete analysis [14] reveals that the maximum number of cells in a single array is simply given by $\exp(E_k/k_BT)$, again requiring excitation energies to be significantly larger than k_BT. The kink energy increases as the system is scaled to smaller sizes. We will see in Section V that we can partition a large problem into smaller problems to surmount this difficulty.

IV. Switching of QCA Arrays

As discussed in the previous section, quantum cellular automata take advantage of the concept of computing with the ground state, which means that the physical ground state

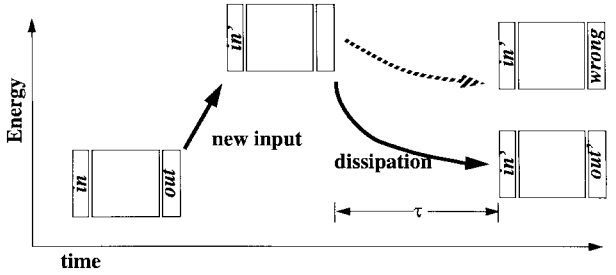


Fig. 8. Schematic representation of a metastable state. Instead of relaxing correctly to the new ground state, a system may be delayed in an excited state due to an inability to tunnel through a kinetic barrier.

of the system is mapped directly to the logical solution of the problem the device is designed to solve. This emphasis on the ground state is one of the strengths of the QCA architecture—the details of the evolution of the system, which may be hard to control, are not essential in getting the computation right. The dynamics of the system are *doing* the computing only in the sense that they move the system to its new ground state. This view of the computational process has also made it appropriate to first study the steady-state behavior of these devices before looking at the dynamic behavior.

The dynamics of the system, however, cannot be completely neglected. The dynamics of the system are relevant for two reasons. The first is that an analysis of the system's dynamics is necessary to quantify the switching speed of QCA arrays. Second, as has been pointed out by Landauer [19], the presence of metastable states could cause a significant delay in the system reaching its new ground state, so the detection of such states is an important goal.

We consider two approaches to switching the array from the solution of one problem to another. The first approach involves switching the input cells suddenly and allowing dissipative coupling to the environment to relax the array to the new solution state. The inputs are kept fixed during this relaxation. The second method involves switching the array gradually by smoothly changing the input states while simultaneously modulating the interdot barriers over the whole array. In this way the array can be switched adiabatically, keeping the system at all times in the instantaneous ground state.

A. Abrupt Switching with Dissipative Coupling to the Environment

If the inputs to a QCA array are switched suddenly the array will be momentarily in some combination of excited states. The excitation energy is provided by the work done on the input cells. The array will then relax to the new ground state by dissipating energy to the environment. This is shown schematically in Fig. 8. Dissipative mechanisms might include phonon and plasmon emission, for example. The detailed description of this dynamic evolution is complicated enormously by the contact between the system and the environment. An accurate theoretical analysis would require knowing all the details of the inelastic channels

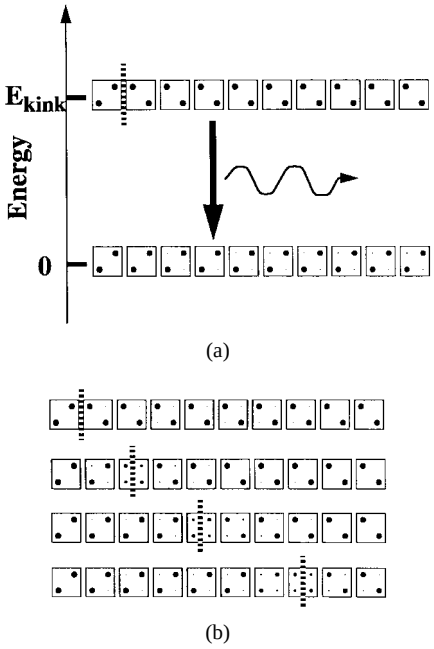


Fig. 9. Schematic representations of the two limits of system-environment interaction. (a) An efficiently dissipative environment yields extrinsic switching times determined by Fermi's Golden Rule. Extrinsic switching of a QCA system. The inelastic relaxation time is much shorter than the elastic switching time, so the system will scatter inelastically to the ground state. (b) An inefficiently dissipative environment demonstrates intrinsic switching and elastic time evolution of the system without interaction with the environment. The elastic switching time is much shorter than the inelastic relaxation time, so the system will exhibit coherent evolution at a constant energy.

available for the dissipative coupling as well as a complete description of the state of the local environment (e.g., the phonon mode occupancy). Assessing the speed at which the system relaxes through these inelastic channels is therefore a matter for experimental determination in particular materials systems and implementations.

While the full dissipative dynamics of the array cannot be solved, we can describe two limits to the coupling between the system and the environment. One limit is when the inelastic relaxation occurs on a much shorter time scale than the elastic (iso-energetic) evolution of the system. This would be the case if the device were operating in an *efficiently dissipative* environment. The coupling to the environment would be very strong and effective at quickly relaxing the system to its ground state.

In this case, the dissipation could be simply characterized by a rate for the transition from the excited state to the new ground state. This rate could be obtained in principle (and perhaps in practice for a simple system) from Fermi's Golden Rule. Such *extrinsic* switching times, which depend critically on the nature of the coupling to the thermal environment, need to be determined experimentally for specific realizations of QCA's. The extrinsic switching of a QCA system is shown schematically in Fig. 9(a).

The other limit of system-environment interaction is when the inelastic coupling to the environment takes place on a time scale much longer than that required for elastic

evolution of the system. In such an *inefficiently dissipative* environment, we can assume that the system is isolated from the reservoir, and model the elastic dynamics using the time-dependent Schrödinger equation, adding in slow dissipation phenomenologically. Simulation of the system using this equation gives information about how signal pulses propagate through the device when energy is not being dissipated to the environment. For this reason, such switching gives information about the *intrinsic* switching speed of the device. Intrinsic switching of a QCA device is shown schematically in Fig. 9(b).

Solutions of the time-dependent Schrödinger equation for the standard semiconductor cell yield an intrinsic switching time of about 2 ps per cell [17]. Intrinsic switching times for the model macro-molecular cells decreases to 0.02 ps per cell.

B. Adiabatic Switching

In many implementations, inelastic processes will likely be very efficient at relaxing the system to its ground state. After all, the usual problem with phase coherent quantum devices is trying to prevent inelastic processes from dominating. Here we are in the enviable position of exploiting the dissipation to accomplish the computing.

Nevertheless, two concerns intrude. One is simply that the inelastic relaxation is uncontrolled and may, in a particular system, be inefficient. The second is that the relaxation may occur in a sense too quickly, and land the system in a metastable state which is a local energy minimum but not the true ground state. The array may then be stuck in a metastable state for a considerable period of time. Fig. 9 illustrates both the desirable relaxation to the true ground state and the undesirable metastable trap.

To avoid these problems we examine here a different mode of switching in which the array always remains in its instantaneous ground state. This approach restores control over all aspects of the switching and eliminates the metastability problem as well.

1) Adiabatic Quantum Mechanics: Adiabatic switching of QCA arrays is based on the quantum version of the adiabatic theorem. The theorem states that if the Hamiltonian of a system undergoes a gradual change from an initial form H^i to a final form H^f, and if a particle starts in the nth nondegenerate eigenstate of the initial Hamiltonian, it will be carried under the time-dependent Schrödinger equation into the nth eigenstate of H^f [20]. In our particular application of this theorem, we will transform the Hamiltonian by lowering the inter-dot barriers within the cell and removing the old input, followed by applying the new input and raising the barriers. If these transitions are carried out gradually, the theorem guarantees that the system, which starts in the ground state of the initial Hamiltonian, will be carried smoothly into the ground state of the new Hamiltonian. In fact, application of the adiabatic theorem at each point in time will guarantee that the system never leaves the instantaneous ground state corresponding to the current inter-dot barriers and driver polarizations.

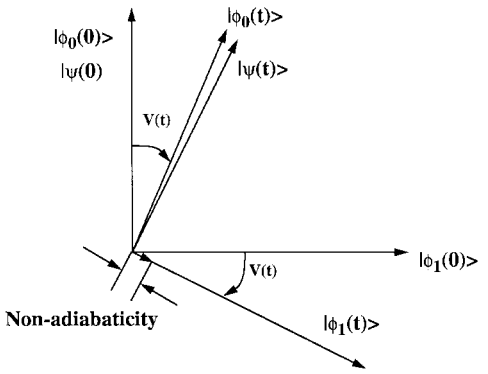


Fig. 10. Schematic representation of the application of the adiabatic theorem to a two-state system. The system starts in the ground state corresponding to the unperturbed Hamiltonian, $\phi_0(0)$. As the Hamiltonian is gradually changed by the application of the potential $V(t)$, the ground state is transformed to $\phi_0(t)$. If the system is changed gradually enough, the state of the system smoothly follows to $\phi_0(t)$. If not, a small amount of nonadiabaticity is introduced as the system state now has a projection on the first excited eigenstate, $\phi_1(t)$.

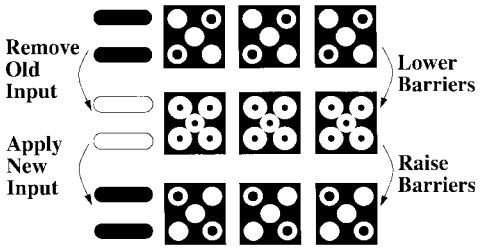


Fig. 11. Schematic of the adiabatic switching of a QCA device. The interdot barriers within each cell are decreased as the old inputs are removed, then the new inputs are applied as the interdot barriers are reasserted. The system smoothly follows its instantaneous ground state, crystallizing in the ground state corresponding to the new input.

Fig. 10 illustrates the process schematically in terms of the Hamiltonian eigen-kets. At the beginning of the transition period ($t = 0$), the particles are in the ground state of the initial Hamiltonian, $|\phi_0(0)\rangle$. As the function $V(t)$ introduces a change to the Hamiltonian, there exists at each time t an instantaneous ground state $|\phi_0(t)\rangle$. The actual time dependent system is represented by $|\psi(t)\rangle$ where $|\psi(t = 0)> = |\phi_0(0)\rangle$. If the transition is gradual enough the system state will track with the instantaneous eigenstate, that is $|\psi(t)\rangle$ will be arbitrarily close to $|\phi_0(t)\rangle$. However, if the transition of the Hamiltonian is not gradual enough, the system will deviate from the instantaneous ground state and have projections on excited states. The extent of this deviation from perfectly adiabatic behavior can be quantified by defining the *nonadiabaticity* $\eta(t)$ as follows:

$$\eta(t) = 1 - |\langle\phi_0(t)|\psi(t)\rangle|^2. \tag{8}$$

In this way η is a measure of that part of the state that is not in the instantaneous ground state of the system. For a switching event which lasts a certain time T, the relevant nonadiabaticity is that which remains when the switching is complete so we also define $\eta_{\text{final}} = \eta(T)$.

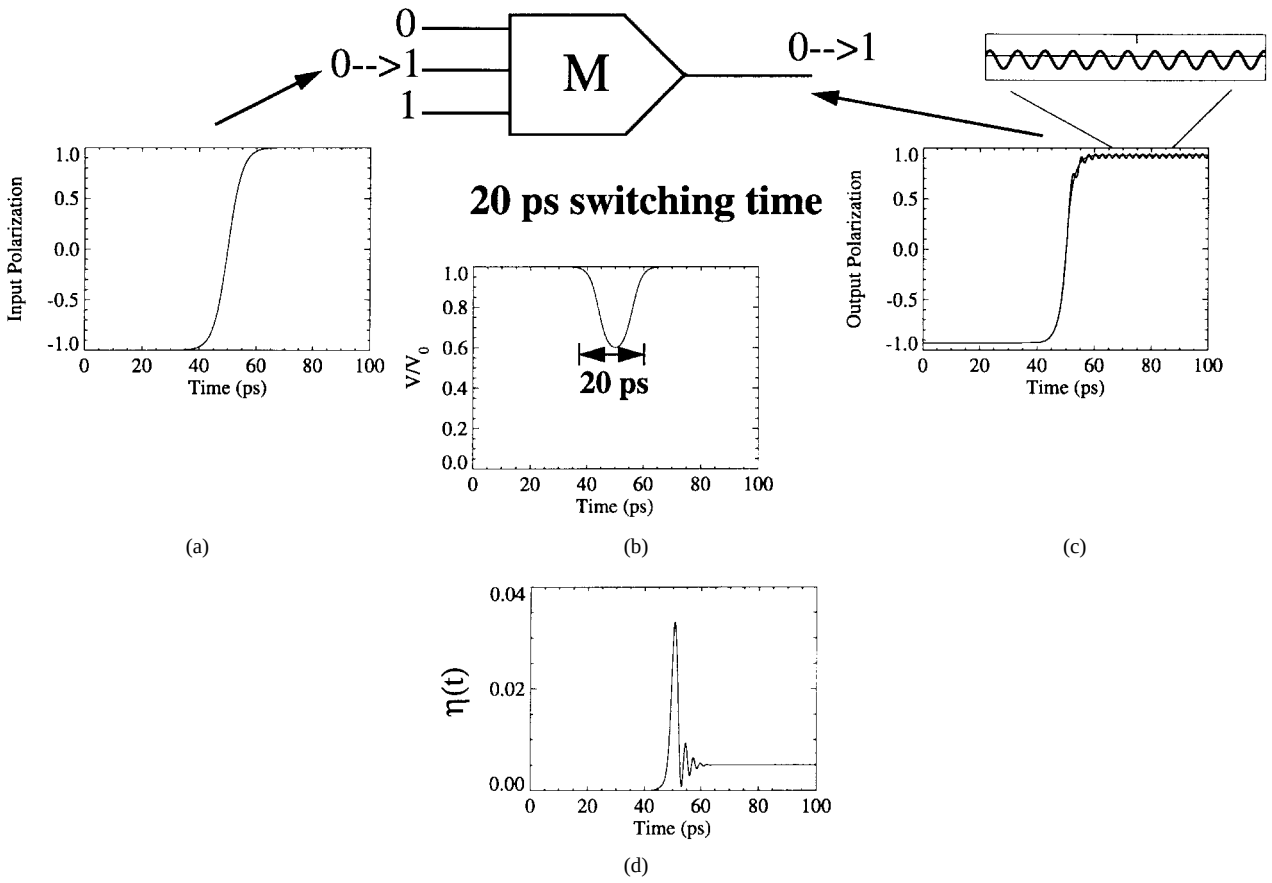


Fig. 12. The adiabatic switching of a single-cell majority logic gate. (a) The interdot tunneling barriers are gradually lowered to one-half their maximum height for a period of 20 ps. (b) The input polarization is switched from -1 to $+1$ over a similar period of time. (c) The output polarization closely follows the instantaneous ground state, but some ringing does develop. (d) $h(t)$ is a measure of the nonadiabaticity introduced into the system. Less than 0.5% ringing is present at the end of the switching cycle.

Fig. 11 shows a schematic diagram of the adiabatic switching of a representative QCA device. The system begins in the ground state appropriate to an old input. The first phase of adiabatic switching includes the lowering of interdot barriers and the gradual removal of the old input. Lowering the interdot barriers reduces the confinement of the electrons on the individual quantum dots, while removal of the old input removes the external boundary condition that was driving the system into one of the two polarization states. The result of each of these actions is to reduce the polarization of the cells in the system, and at the end of this first phase of switching, the cells exhibit little or no polarization. The two-electron wavefunctions have delocalized across the cell.

The second phase of switching includes raising the interdot barriers while the new inputs are being applied. The increased interdot barriers cause the cells to repolarize into well defined bistable states and the cells "crystallize" into the ground state corresponding to the new inputs. Raising the interdot barriers localizes the electrons on the individual quantum dots, while applying the new input drives the system into one of the two polarization states. Note that across one adiabatic cycle no work is done by the interdot barrier potential being raised, lowered and then raised. Moreover, we do not require separate connections to each cell; only one additional "clock" potential needs to be present across the entire array.

2) Adiabatic Switching of a Majority Gate: As a concrete example of adiabatic switching of a QCA device, we consider the switching of a single-cell majority logic gate similar to the one shown schematically in Fig. 3(d). The three drivers are represented by pairs of fixed charges on the top, bottom, and left of the device cell. The top and bottom drivers are set in opposite polarization directions, so the state of the device cell is determined by the state of the left driver. The left driver, the "tie-breaker" in this case, is switched from the zero state to the one state, and a corresponding change in the output polarization is seen. The charges on the driver electrodes correspond to those present in neighboring cells of a QCA array, but they could also be the electrodes used to apply the inputs to the edge of the device. Sensing circuitry will be necessary to measure the switching of the output polarization.

During the adiabatic switching cycle, the interdot barriers within the device cell are lowered and then raised. Fig. 12(a) shows the height of the interdot barrier used for our first calculation. The barriers are lowered to approximately one-half of their original height and then raised back up

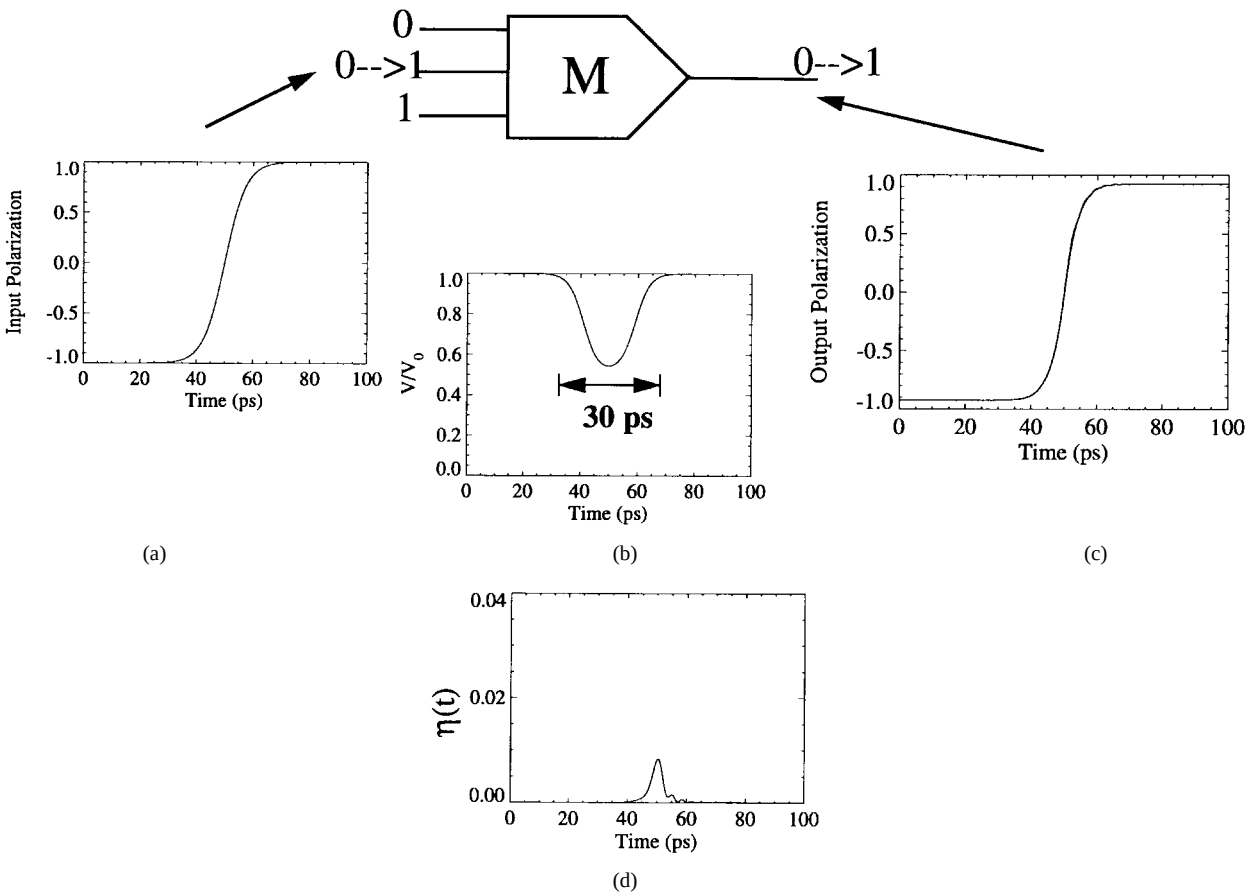


Fig. 13. Adiabatic switching of the device of Fig. 12 with slower transitions. The calculation to generate this figure was identical to that for Fig. 12, except that a longer transition time is used for the tunneling barriers and input polarization. This longer tunneling time leads to decreased ringing in (c) and a very low amount of nonadiabaticity in (d).

to V_0. The shape of the barrier height profile is generated using Gaussian functions, which provide smooth transitions between the high and low barrier states. In this simulation, the total switching time is shown to be approximately 20 ps (ten times the intrinsic switching time).

Fig. 12(b) shows the corresponding function used for P_{driver}. This function switches from -1 to $+1$ using a Gaussian-derived transition function while the interdot barriers within the cell are low. Since the other two inputs of the majority logic gate are conflicted, the state of this third input determines the correct output for the device. In this way, the output of the device should exactly follow the logical state of the third input.

Fig. 12(c) shows the output polarization and the instantaneous ground state as the device is switched. The output is calculated from a direct solution of the time-dependent Schrödinger equation. If adiabaticity were assured, we could simply solve for the ground state at each time; however, we are looking here for deviations from adiabaticity so the full time-dependent problem must be solved. It can be seen that the device switches from the 0 state to the 1 state as required, but some slight ringing in the output polarization has been introduced indicating a small admixture of excited states. The nonadiabaticity shown if Fig. 12(d), though small, indicates that the device is no longer completely in the ground state. Fig. 13 shows the comparable switching event with a slightly longer switching time of 30 ps. In this case the nonadiabaticity is greatly reduced. The adiabatic switching time of macro-molecular cells (see Section III-D above) are reduced to 0.2–0.3 ps.

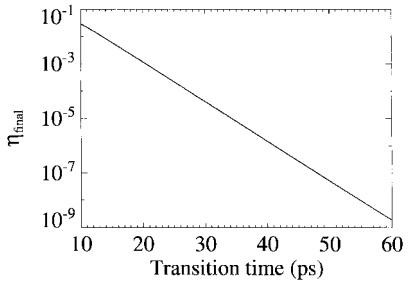


Fig. 14. Scaling of nonadiabaticity with the transition time. The switching functions are based on Gaussian curves, as in Figs. 12 and 13. The linear shape of this semi-log curve indicates an exponential decrease in nonadiabaticity with transition time.

Fig. 14 shows the dependence of the residual nonadiabaticity η_{final} on the switching time for a QCA majority gate. The nonadiabaticity is reduced exponentially as the switching time is increased. This is extremely significant since it means that unwanted ringing can be easily and effectively reduced. Moreover, in a real system small exci-

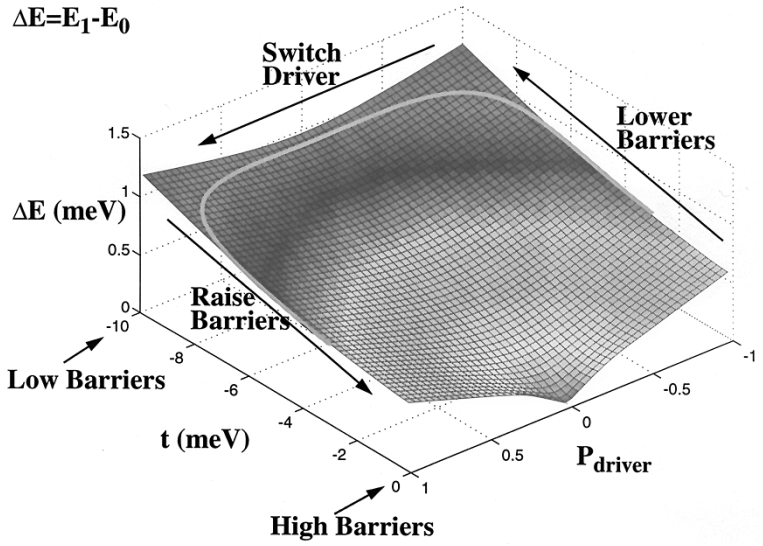


Fig. 15. The energy splitting between the ground state and the first excited state of a QCA cell as a function of P_{driver} and tunneling coefficient. An adiabatically switched system follows the curve traced on the surface. As barriers are lowered, the driver is switched, and the barriers are reasserted, the system remains in areas with relatively high splitting between the ground state and the excited states. One must be careful to avoid the dip in energy splitting where the barriers are high and the driver is removed, as it would introduce significant nonadiabaticity or require extremely long switching times.

tations would undoubtedly be damped by inelastic processes we are neglecting in this treatment.

3) Adiabatic Path: Examination of the proof of the adiabatic theorem shows that the term "gradual transition" really means a transition time longer than the time corresponding to energy splitting between the ground state and the excited states of the system, i.e., $T \gg \hbar/\Delta E$. For this reason, it is important that this energy splitting be maximized, since passing through a region where there is very little splitting between the ground state and the excited states would require very slow switching times.

Fig. 15 shows a plot of the energy splitting between the ground state and the first excited state of a QCA majority gate undergoing adiabatic switching. The line drawn on the surface of the plot indicates the path through this space that is taken by the system as it is switched according to the prescription above. As the switching proceeds—the barriers are lowered, the driver is switched, then the barriers are reasserted—a path is traced along that energy surface. In order to enable maximum switching speed, it is important that the system remain in areas of the surface where the energy splitting is high.

As seen in this figure, the system should stay away from the point where the driver polarization is zero and the barriers are high. At this point, there is very little energy splitting between the ground state and the first excited state, and the switching would either introduce significant nonadiabaticity, or it would have to be switched extraordinarily slowly. This is the reason the interdot barriers are lowered during the switching.

4) Scaling with System Size: Clearly for larger systems the adiabatic switching must proceed more slowly than for smaller systems. We now consider how the switching time and nonadiabaticity vary with the number of cells in a linear array. Assessing this scaling requires several steps and involves quite a large number of solutions of the time-dependent Schrödinger equation.

The first of these calculations is illustrated in Fig. 16(a), which shows a schematic representation of the adiabatic switching of a line of five cells and a driver cell. At point **{a**, the system has not yet begun to switch. The barriers are high and the driver is fully polarized. The other cells are in the ground state corresponding to the fully polarized driver. At point **{b**, the driver has been relaxed and the interdot barriers have been lowered. These two actions have led the other cells in the line to have little or no polarization. In state **{c**, the driver has finished switching and the interdot barriers have begun to be reasserted. The other cells still don't have a great deal of polarization, but are beginning to polarize. State **{d** shows the device after the full switching event has occurred; the driver is fully polarized, the interdot barriers are back to their original height, and the nondriver cells in the line are in the new ground state corresponding to the new driver polarization.

Fig. 16(b) shows a plot of the calculated polarization of the last cell in the line as a function of time. This cell, along with all the other cells in the line, is undergoing the switching process just outlined above. The letters next to the curves of this figure correspond to the four states shown in Fig. 16(a), and show the approximate positioning of each of those states. The two curves correspond to the time-dependent state of the system, including nonadiabaticity, and the instantaneous ground state of the system. If the array were being switched perfectly adiabatically, these two curves would be identical, but the finite transition time used to switch this device leads to a certain amount of

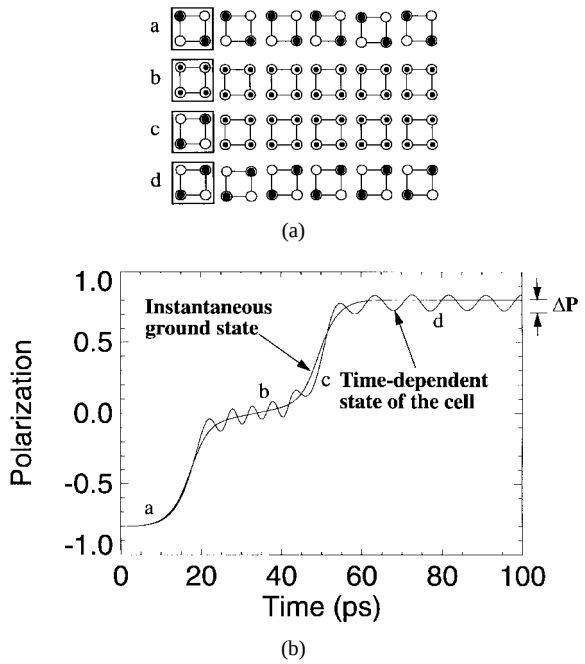


Fig. 16. The nonadiabatic ringing introduced in a line of QCA cells during adiabatic switching. (a) Schematic representation of four states of a line during the process of a switching cycle. (b) The cell polarization and the instantaneous ground state polarization of the last cell in a line of five cells being adiabatically switched. Nonadiabatic ringing introduces a small amount of oscillation to the cell polarization. This is characterized by ΔP, the maximum absolute variation between the cell polarization and the ground state polarization after the switching cycle is complete. ΔP for the last cell in a line is always the largest, so we will take that value to represent the nonadiabatic ringing introduced into the system.

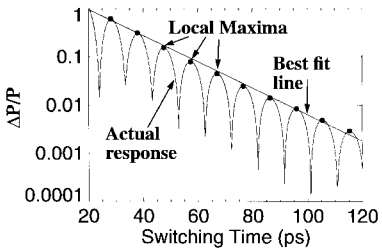


Fig. 17. Variation of nonadiabatic ringing in the five-cell line with total switching time. Nonadiabatic ringing is measured by $\Delta P/P$ of the last cell in the line. The actual response of the line showed a series of local minima and maxima. The local maxima were identified, and a best fit envelope was drawn through the lines. This envelope provides an upper limit on the error introduced to the line at a particular switching speed.

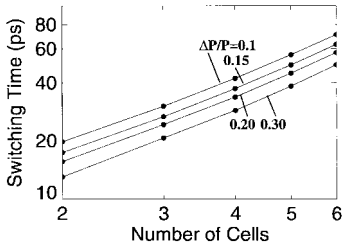


Fig. 18. Log scale plot of the minimum allowable switching time for a line of cells as a function of the number of cells and the maximum allowable nonadiabatic ringing. The switching time dependence upon the number of cells appears to follow a power law, and a bet-fit line through the top curve gives a slope of 1.16. Thus the minimum switching time varies in an almost linear way with the number of cells in a line.

nonadiabaticity, which is shown in this figure by the small oscillation of the time-dependent state of the system.

For device operation we are not really concerned with nonadiabaticity of the internal cells, what we require is a clear output signal. We therefore choose to focus on the oscillation in the last cell as a measure of the relevant nonadiabaticity. We define the *nonadiabatic ringing* ΔP to be the absolute value of the maximum difference between the cell state and the ground state after the switching is complete. In order to account for variation of the ground state polarization P between devices, we will normalize ΔP by dividing by P. Thus in a calculation very similar to that previously shown for the single-cell majority logic gate, we have defined a system with a variable number of cells and identified $\Delta P/P$ as a measure of the nonadiabaticity present in the array.

Once the nonadiabaticity induced in a particular line has been calculated for a particular switching time, the transition time can then be varied to see how the nonadiabaticity decreases as the transition becomes more gradual. In Fig. 17, we see the variation of $\Delta P/P$ for the last cell in a line of five cells as the total transition time is increased. While overall there is an exponential decrease, there is also surprising structure. The response shows a series of very sharp local minima, where the nonadiabatic ringing essentially vanishes, occurring at regularly spaced intervals of approximately 10 ps. The reason for the rapid decreases in nonadiabaticity is because we are driving the device at or near a multiple of the natural frequency of its abrupt switching speed. Recall that abruptly switched QCA cells propagate information at a rate of approximately 2 ps per cell. For a line of five cells, this would indicate an abrupt switching speed of 10 ps. Since an "adiabatically" switched system with a short switching time is essentially the same as an abruptly switched device, we see local decreases in error at the frequencies corresponding to timing the transition just right to abruptly switch the device with little or no error. We ignore these special "super-switching" points and focus on the overall trend. We locate the local maxima of this series of peaks and valleys and include only their values in the calculation of the nonadiabaticity. We draw a best-fit envelope through these points and use this envelope as an upper limit on the nonadiabaticity ringing.

Knowing the nonadiabaticity introduced at a particular transition time, we can now calculate the minimum allowable transition time to meet varying levels of allowable nonadiabaticity ringing. This process is then repeated for various numbers of cells in the line. The results are plotted on a log-log scale in Fig. 18.

This figure shows four lines, each of which represents a different level of allowable nonadiabatic ringing. As the number of cells increases, the minimum time required to switch a device at a particular error level also increases. Since this dependence is approximately linear in a log-scale plot, the dependence of the switching time on the number

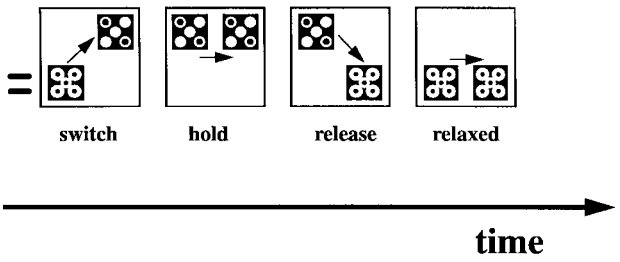


Fig. 19. The four stages of adiabatic pipelining. Within each box, the cell on the left represents the state of the cells at the beginning of the cycle, while the cell on the right represents the state of the cell at the end of the cycle. The relative height of the cells within the box represents the height of their interdot barriers; lower cells have low barriers, while higher cells have higher barriers. As the cell passes through these four stages, the barriers are raised, then held high, then lowered, then held low.

of cells in an array appears to follow a power law. When a best-fit line is calculated for the top curve, it is found that the slope of this line is 1.16. Therefore, the minimum switching time dependence on the number of cells in an array is

$$T \propto N_{\text{cells}}^{1.16}. \tag{9}$$

Intuitively, one would expect a linear relation—it takes twice as long to adiabatically switch twice as many cells. That we find an exponent slightly larger than one may in fact be due to the artifice of fitting the maxima in Fig. 17.

It should be noted that the nonadiabatic ringing in the outputs does not necessarily constitute an error in the computed result unless P actually changes sign. As long as the sign of the polarization can be measured, the correct result can be obtained. A more detailed treatment is really required to establish appropriate noise margin analysis for these devices.

This calculation for a linear array of cells may be simply a worst-case limit for more general arrays. In the line, switching the input causes all the cells in the array to switch. Analysis of other topologies indicates that cells whose state does not change when a given input is switched do not contribute significantly to the array switching time. A complete understanding of the relationship of array topology to the switching time is not yet in hand.

V. Pipelining Adiabatically Switched Circuits

In the adiabatic switching approach described in the previous section, it was always assumed that the interdot potential barrier was being modulated simultaneously for all cells in the array. From the point of view of fabrication complexity, this is an important feature. It permits one conductor, typically one gate electrode, to control the barriers of all cells. If each cell had to be separately timed and controlled, the wiring problem introduced could easily overwhelm the simplification won by the inherent local interconnectivity of the QCA architecture itself.

We can gain significant advantage, however, by relaxing this requirement slightly. If we subdivide an array of cells into subarrays, we can partition the computational problem and gain the advantages of multi-phase clocking and pipelining. For each sub-array a single potential (or gate) modulates the inter-dot barriers in all the cells. This enables us to use one sub-array to perform a certain calculation, then freeze its state by raising the inter-dot barriers and use the output of that array as the input to a successor array. During the calculation phase, the successor array is kept in the unpolarized state so that it does not influence the calculation.

Fig. 19 shows a schematic diagram of the four stages required for an adiabatic pipelining cycle. Each of the boxes represents the state of a subsystem of several cells. Each cell within the subsystem has the same gate controlling inter-dot barriers. Within each box, the schematic cell on the left represents the state of the cells at the beginning of this clock phase, while the cell on the right represents the state of the cells at the end of the clock phase. We emphasize that although only two cells are shown in each subarray, they are meant to be representative of a larger number of cells.

During the first phase, called the *switch* phase, the cells begin unpolarized and with low barriers, but the barriers are raised during this phase and the cells become polarized according to the state of their driver. This is the clock phase during which actual computation occurs. By the end of this clock phase, the barriers are high enough to suppress any tunneling and the cell states are essentially fixed. During the next clock phase, the *hold* phase, the barriers are held at this high value so the outputs of the subarray can be used as inputs to the next stage. Next, in the *release* phase, the barriers are lowered and the cells are allowed to relax to an unpolarized state. During the fourth clock phase, the *relaxed* phase, the cell barriers remain lowered, keeping the cells in an unpolarized neutral state. After this fourth phase, the subsystem will return to the first clock phase and repolarize.

To illustrate the interaction between these phases in the simplest possible system, we consider an adiabatically pipelined wire. A schematic representation is shown in Fig. 20. The figure shows six subsystems along the length of the wire, each of which is assigned an adiabatic clock. Since there are six of these stages, the fifth and sixth stage repeat the clock signals of the first and second stage. The current state of this wire shows a switching event in the second and sixth stages, which are switching to 1 and 0, respectively. In such a pipelined wire, it is possible to be transmitting more than one bit of information at a time. A wire such as this exhibits minimum energy communication between two points [21].

Fig. 21 shows the general scheme for an adiabatic pipelining system composed of several sub-arrays. Different time steps (clock phases) are separated vertically in the figure, while subarrays that are physically located next to each other are shown horizontally. In the first time step, all of the cells begin unpolarized, but the first stage carries out a "switch" and polarizes according to the driver polarization applied at the edge of the device. In the second time step, the first stage is held fixed while its neighbor is allowed

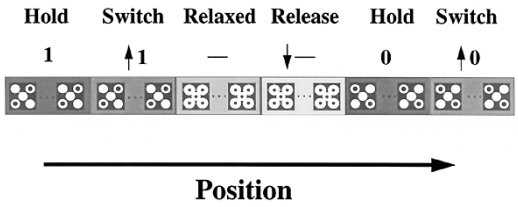


Fig. 20. Schematic representation of a long line of adiabatically clocked cells. Each subsystem contains many cells, which are represented schematically by the two end cells. Note that this adiabatically switched wire currently contains two bits of information, since the pipeline is six stages long.

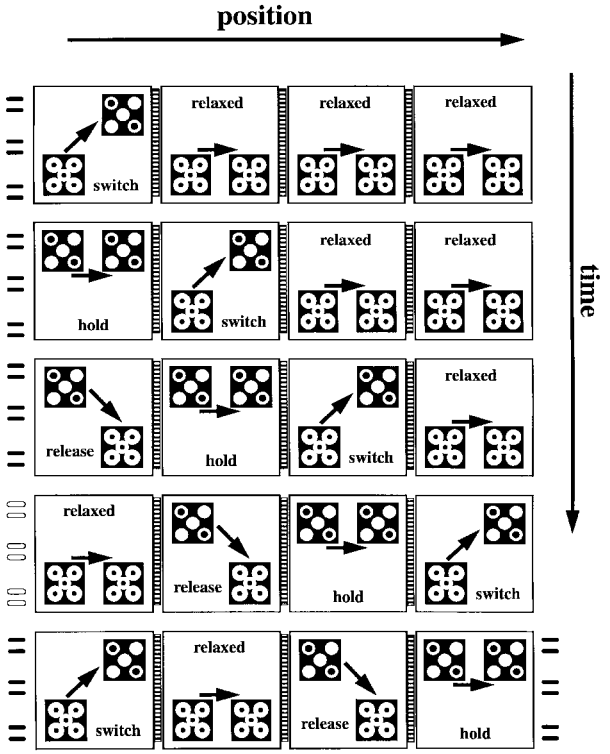


Fig. 21. Adiabatic pipelining of QCA devices. Several clocked subsystems are used to drive one another and maximize throughput of the device, since more than one calculation can be "in the pipeline" of a single device at a given time. In addition, subsystems in the "switch" phase have one "held" neighbor and one "relaxed" neighbor, so there is clear unidirectionality of the calculation path.

to switch according to the first stage's polarization. In the third time step, the first stage, which has already propagated its information to the next stage, is allowed to be released. The second stage is held fixed, while the third reacts to its polarization. In the fourth time step, the first stage is relaxed, the second is released, the third is held, and the fourth is switched. The cycle then repeats itself as a new input from the edge is clocked into the switching of the first stage. If several of the outputs from the last sub-array are input to the first, a finite state machine is formed, showing in principle that general-purpose computing is possible.

A *caveat* regarding the role of dissipation in this scheme should be added here. As Landauer has shown, there is a connection between logical reversibility and physically reversible processes. During the *relax* phase of each sub-array's cycle, the physical system can evolve strictly

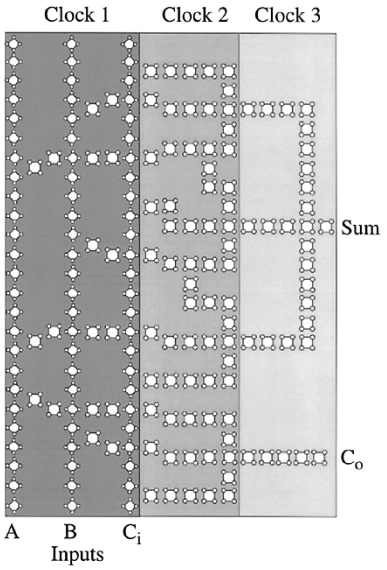


Fig. 22. Layout of an adiabatically pipelined QCA full adder circuit. Information is applied at the three inputs and distributed to the majority logic gates in clock phase one. Phase two carries out the first level of majority logic, while phase three carries out the second level of majority logic.

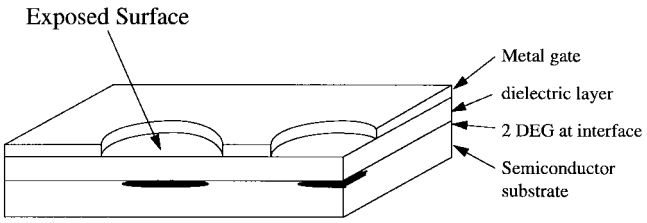


Fig. 23. A possible physical implementation of a QCA cell. The metal top gate introduces electric fields in the substrate to deplete electrons in the 2–D electron gas formed at the junction of the dielectric layer and the substrate. Quantum dots form at locations where the metal gate has been removed to leave an exposed surface. By patterning the areas of exposed surface, it is possible to create any layout of quantum dots needed. This structure could be implemented using GaAs or a combination of silicon and silicon dioxide.

adiabatically only if the computation it is carrying out is logically reversible. The logical reversibility means there will be enough output cells on the right to keep the system from relaxing in an uncontrolled way. This is strictly the case in the wire shown in Fig. 19; it is for exactly this reason that it can be considered an example of Landauer's principle of minimum energy communication. In a more general situation one must either design the array to be logically reversible, or simply allow a slight dissipation to settle the state into the new ground state. (Our adiabatic clocking technique can be regarded as simply an application in a concrete physical system of the general ideas worked out by Landauer more than two decades ago [22].)

A concrete example of an adiabatic pipeline is shown schematically in Fig. 22. This schematic diagram is based on the design of the QCA full adder circuit of Fig. 5. Three inputs are applied to the device at the bottom, and this information is used to drive the three parallel lines and connecting horizontal lines labeled as clock 1 in this figure.

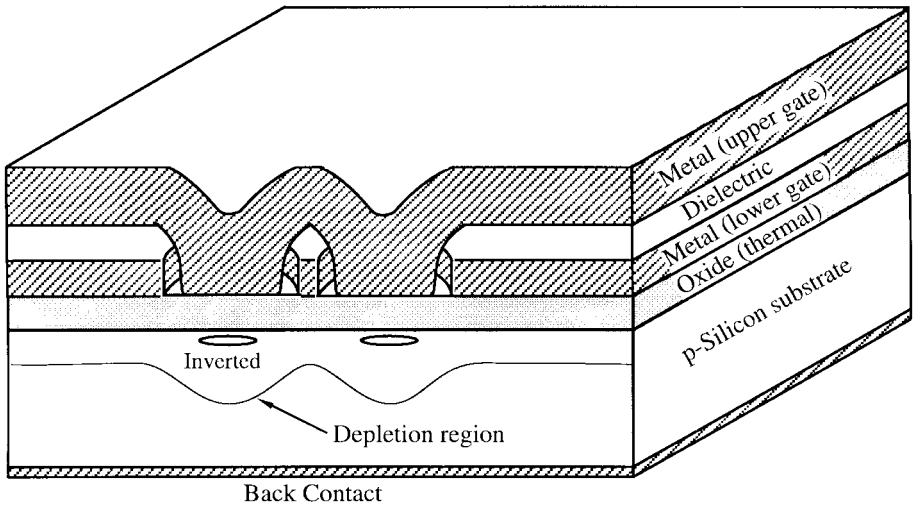


Fig. 24. A realization of QCA cells in a silicon system. Here, two top metal gates are used to control the occupancy of electrons in the p-type silicon substrate. The lower gate is used to deplete holes near the surface of the substrate at all locations where quantum dots are not desired. The upper gate then inverts the p-type substrate at particular locations, leading to the creation of quantum dots. This combination of depletion and inversion provides excellent control over the size and position of the dots.

Thus, the first clock phase accepts the three inputs from an outside source and distributes the signals to the various majority logic gates needed to carry out the calculation. Clock 2 drives the majority logic gates that carry out the first level of logical calculation, while clock 3 controls the final level of logic. This demonstrates how a complex calculation can be broken up into several logical stages, and a separate clock signal can be used to drive each of those stages.

The adiabatic pipelining scheme has several benefits. The most obvious benefit is that the clocking cycles of the cells are interlaced so that as soon as information is no longer necessary for further calculations, it is released to free up room for new information. This allows the device to be in the process of carrying out several calculations at once, especially if the pipeline is long. Such simultaneous calculation stages maximize the throughput of each total system. A second benefit of this system is that the number of cells in each subarray can be kept well below the thermodynamic limit discussed in Section III-D. Finally, this clocked approach clearly demonstrates that, at least from an architectural standpoint, general purpose computing with QCA arrays is feasible.

VI. Possible QCA Implementations

Much of the work on quantum dots has focused on dots in semiconductors. One well-developed technique uses metal top-gates to pattern the potential energy surface for electrons already confined in a 2-D electron gas (2DEG) at an interface. The interface can be a III–V heterojunction or a Si–SiO_2 interface. This is shown schematically in Fig. 23. Even sharper dot profiles are possible using a double-gate configuration as shown in Fig. 24. This double top gate also enables adiabatic switching by providing separate control of inter-dot barriers. Using these approaches dots can be made small enough to hold a single electron. It should be noted that it is not strictly necessary to have zero or one electron in each dot for the QCA concepts to work. If the dots have many electrons they can be treated as metallic puddles with a possible extra electron. We have shown that metallic systems, which can be viewed as tunnel-junction capacitors, have essentially identical properties [12].

Advances in scanning tunneling lithography hold the promise of fabrication down to the angstrom level [23]. Much work remains to be done to transfer currently produced patterns into a material system which could sustain quantum dots. All lithographic approaches need to deal with the problem of stray charges. This is very much a materials problem but it is one that vexes any attempt to approach the ultimate limits of few-electron computing. The self-correcting nature of QCA wires mitigates but does not eliminate this problem [13].

One candidate for a true molecular realization using metal cluster carboxylate molecules [24] is shown in Fig. 25. While molecular implementations present many serious challenges, particularly in input and output, they have the advantages of perfectly uniform cell size and enormous numbers of cells on which to experiment. Electrochemical and optical experiments are possible because of the large arrays that can be synthesized.

Finally, it can be seen from the description of the QCA paradigm at the end of Section II that the concept is not inherently restricted to electrostatic coupling—any physical coupling which yields sufficient nonlinearity in the cell—cell response function would do. In particular, it may be possible to implement the system with small magnetic arrays [25].

VII. Conclusions

Quantum dot fabrication technology deserves to be pursued because it is possible to conceive of architectures which will enable coupled quantum dots to perform real

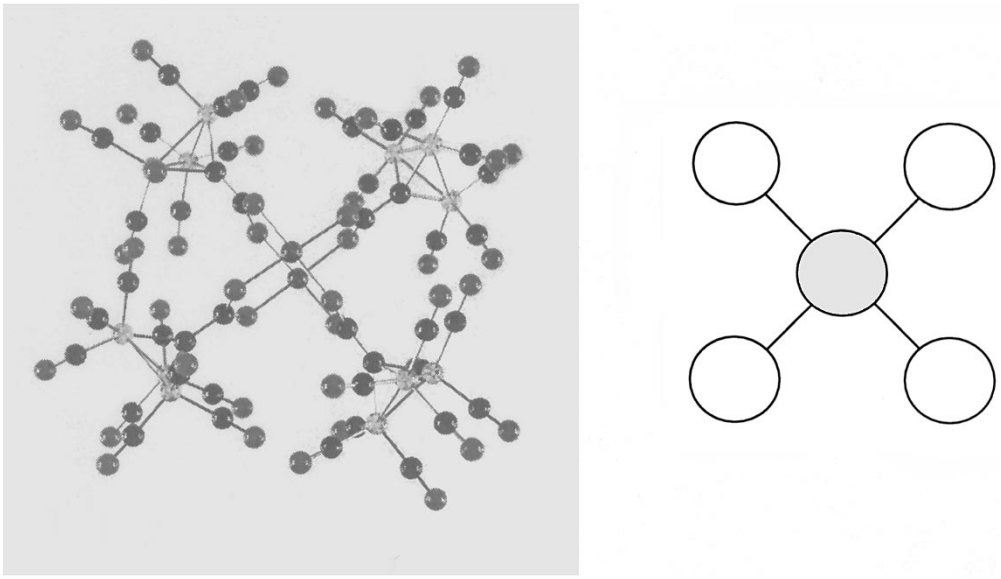


Fig. 25. Possible molecular implementation of the QCA concept. The molecule $M_2\{(CO)_9(Co_3CCO_2\}^4$, where M = Mo, Mn, Fe, Co, or Cu, can be synthesized and exhibits stable charge states. Each of the four Co clusters plays the role of a quantum dot in the QCA cell.

computing. We have outlined the QCA approach in a simplified intuitive way and discussed the more rigorous quantum mechanical treatment. We have focused particular attention on adiabatic switching and pipelining because these seem to offer the most flexible use of QCA circuits.

It deserves to be emphasized that the key first step is constructing a device, the QCA cell, which does not reproduce transistor action, but rather exploits the inherent physics of inter-dot coupling to represent binary information in a robust way.

Potential speeds of QCA devices are impressive, but perhaps their real power lies in the parallelism possible with large arrays. Speed and packing densities increase enormously as devices are scaled down. The limit of scaling appears to be set by the size of atoms, a limit unlikely to be circumvented.

Fabrication of QCA devices presents a great challenge, one which is currently being undertaken by several laboratories using several approaches. This is a long-range undertaking. An immediate benefit is that having a clearer architectural goal provides a focus and, to some extent offers a benchmark, for nanofabrication techniques of varying types.

Other investigators have been extending the theoretical analysis of QCA arrays. Tanamoto *et al*. [26] have proposed alternative ways of assembling QCA cells into useful devices. Lusth and Jackson [27]–[29] have applied graph theoretic analysis to QCA design. Chen and Porod [30] have developed sophisticated finite element models for gate-depleted quantum dots in semiconductors that can relate dot occupancy to particular bias conditions. Fountain and coworkers [31] have been matching the QCA approach with massively parallel processing schemes which require only simple computational elements at each node.

Clearly the current state of the concept will be further developed and refined both in the light of experimental results and because architectural and circuit-theoretical implications are only beginning to come into view. One point obvious from the layout of Fig. 5 is that in the current approach, QCA cells seem to be underused. Despite their inherent processing power, the vast majority of cells are being employed in wires. The self-correcting nature of the wire so formed is a great advantage, of course. Nevertheless it is suggestive that a more radical departure from Boolean logic circuits might harness this computational power more naturally. Recent work framing the QCA paradigm in the language of Cellular Neural Networks [16] represents an initial step in this direction. Further efforts to work out implications and constraints at higher levels of design will require joint efforts with circuit theorist, computer architects, and computer scientists.

QCA switching simulations can be seen with a Java-enabled browser at http://www.nd.edu/∼lent/QCA home.html.

ACKNOWLEDGMENT

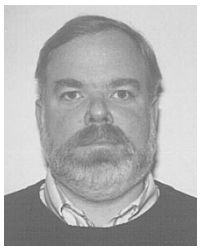
The authors gratefully acknowledge collaboration with W. Porod, G. H. Bernstein, J. Merz, and G. Snider of the Notre Dame Nanoelectronic Group. They also have been aided by conversations with R.-W. Liu and Y.-F. Huang of Notre Dame's systems group.

REFERENCES

[1] R. C. Ashoori, H. L. Stormer, J. S. Weiner, L. N. Pfeiffer, K. W. Baldwin, and K. W. West, "N-electron ground state energies of a quantum dot in a magnetic field," *Phys. Rev. Lett.*, vol. 71, pp. 613–616, 1993.

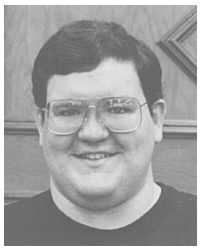
[2] B. Meurer, D. Heitmann, and K. Ploog, "Single-electron charging of quantum-dot atoms," *Phys. Rev. Lett.*, vol. 68, pp. 1371–1374, 1992.

[3] F. R. Waugh, M. J. Berry, D. J. Mar, R. M. Westervelt, K. L. Campman, and A. C. Gossard, "Single-electron charging in double and triple quantum dots with tunable coupling," *Phys. Rev. Lett.*, vol. 75, pp. 705–708, July, 1995.
[4] F. Hofmann, T. Heinzel, D. A. Wharam, J. P. Kotthaus, G. Bohm, W. Klein, G. Trankle, and G. Weimann, "Single electron switching in a parallel quantum dot," *Phys. Rev. B*, vol. 51, pp. 13872–13875, May, 1995.
[5] R. H. Blick, R. J. Haug, and K. Eberl, "Single electron tunneling through a double quantum dot: The artificial molecule," *Phys. Rev. B*, vol. 53, p. 7899, 1996.
[6] L. J. Geerligs, C. J. P. M. Harmans, and L. P. Kouwenhoven, Eds., "The physics of few-electron nanostructures," *Physica B*, vol. 189, 1993.
[7] C. S. Lent, P. D. Tougaw, and W. Porod, "Bistable saturation in coupled quantum dots for quantum cellular automata," *Appl. Phys. Lett.*, vol. 62, pp. 714–715, Feb. 1993.
[8] C. S. Lent, P. D. Tougaw, W. Porod, and Gary H. Bernstein, "Quantum cellular automata," *Nanotechnology*, vol. 4, pp. 49–57, 1993.
[9] P. D. Tougaw, C. S. Lent, and W. Porod, "Bistable saturation in coupled quantum-dot cells," *J. Appl. Phys.*, vol. 74, pp. 3558–3566, Sept. 1993.
[10] C. S. Lent and P. D. Tougaw, "Lines of interacting quantum-dot cells: A binary wire," *J. Appl. Phys.*, vol. 74, pp. 6227–6233, Nov. 1993.
[11] P. D. Tougaw and C. S. Lent, "Logical devices implemented using quantum cellular automata," *J. Appl. Phys.*, vol. 75, pp. 1818–1825, Feb. 1994.
[12] C. S. Lent and P. D. Tougaw, "Bistable saturation due to single electron charging in rings of tunnel junctions," *J. Appl. Phys.*, vol. 75, pp. 4077–4080, Apr. 1994.
[13] P. D. Tougaw and C. S. Lent, "Effect of stray charge on quantum cellular automata," *Jap. J. Appl. Phys.*, vol. 34, pp. 4373–4375.
[14] C. S. Lent, P. D. Tougaw, and W. Porod, "Quantum cellular automata: The physics of computing with quantum dot molecules," *PhysComp '94: Proc. Workshop Phys. Comp.*, IEEE Computer Society Press, 1994.
[15] P. D. Tougaw and C. S. Lent, "Dynamic behavior of coupled quantum dot cells," in *Proc. 3rd Int. Conf. Computational Elec.*, S. Goodnick, Ed., May 1994.
[16] G. Toth, C. S. Lent, P. D. Tougaw, Y. Brazhnik, W. Weng, W. Porod, R.-W. Liu, and Y.-F. Huang, "Quantum cellular neural networks," *Superlattices and Microstructures*, vol. 20, pp. 473–478, 1996.
[17] C. S. Lent and P. D. Tougaw, "Dynamics of quantum cellular automata," *J. Appl. Phys.*, vol. 80, pp. 4722–4736, Oct. 1996.
[18] Noninvasive probing of single-electron charging in a semiconductor quantum dot has recently been reported by M. Field, C. G. Smith, M. Pepper, J. E. F. Frost, G. A. C. Jones, and D. G. Hasko, "Measurements of coulomb blockade with a noninvasive voltage probe," *Phys. Rev. Lett.*, vol. 70, pp. 1311–1314, Mar. 1993.
[19] R. Landauer, *Ultimate Limits of Fabrication and Measurement*, M. E. Welland, Ed. Dordrecht: Kluwer, 1994.
[20] D. J. Griffiths, *Introduction to Quantum Mechanics.* Englewood Cliffs, NJ: Prentice-Hall, 1994.
[21] R. Landauer, "Minimal energy requirements in communication," *Sci.*, vol. 272, p. 1914, 1996.
[22] R. W. Keyes and R. Landauer, "Minimal energy dissipation in logic," *IBM J. R&D*, vol. 14, pp. 152–157, Mar. 1970.
[23] J. W. Lyding, T.-C. Shen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, "Nanoscale patterning and oxidation of H-passivated Si(100)-2x1 surfaces with an ultrahigh vacuum scanning tunneling microscope," *Appl. Phys. Lett.*, vol. 64, p. 2010, 1994.
[24] W. Cen, K. J. Haller, and T. P. Fehlner, "On the role of PES data in the identification of metal–metal charge transfer bands in clusters of clusters," *J. Electron Spectroscopy*, vol. 66, pp. 29–36, 1993.
[25] D. H. Pearson and R. J. Tonucci, "Nanochannel glass replica membranes," *Sci.*, vol. 270, p. 68, 1985.
[26] T. Tanamoto, R. Katoh, and Y. Naruse, "A novel quantum cellular automata logic with loop structures," *Jap. J. Appl. Phys.*, vol. 33, pp. 1502–1505, Oct. 1994.
[27] J. C. Lusth and D. J. Jackson, "Graph theoretic approach to quantum cellular design and analysis" *J. Appl. Phys.*, vol. 79, no. 4, p. 2097, Feb. 1996.
[28] J. C. Lusth, "An introduction to optimal colorings of QCA graphs," in *Proc. 34th ACM SE Regional Conf.*, Apr. 1996.
[29] J. C. Lusth, B. Dixon, D. J. Jackson, and S. L. Burkett, "Quantum-dot cellular automata and the problem of unbalanced logic gates," to be published in *Int. Conf. Quan. Circ. and Dev*,
[30] M. Chen and W. Porod, "Design of gate-confined quantum dot structures in the few-electron regime," *J. Appl. Phys.*, vol. 78, pp. 1050–1057, 1995.
[31] T. Fountain, unpublished.



Craig S. Lent received the A.B. degree in physics from the University of California at Berkeley and the Ph.D. degree in solid-state physics from the Univerity of Minnesota, Minneapolis-St. Paul.

In 1986 he joined the faculty at the University of Notre Dame, Notre Dame, IN, where he is now Professor of Electrical Engineering. His research interests include RHEED oscillations, deep level defects in semiconductors, and semiconductor quantum devices.



P. Douglas Tougaw received the B.S.E.E. degree from Rose-Hulman Institute of Technology, Terre Haute, IN, and the Ph.D. degree in electrical engineering from the University of Notre Dame, Notre Dame, IN.

He is Assistant Professor of Electrical and Computer Engineering at Valparaiso University, Valparaiso, IN. His research interests are in the area of quantum devices and the physics of semiconductor devices.