CrossNets

High-Performance Neuromorphic Architectures for CMOL Circuits

KONSTANTIN LIKHAREV, ANDREAS MAYR, IBRAHIM MUCKRA, AND ÖZGÜR TÜREL
Stony Brook University, Stony Brook, New York, USA

ABSTRACT: The exponential, Moore’s Law, progress of electronics may be continued beyond the 10-nm frontier if the currently dominant CMOS technology is replaced by hybrid CMOL circuits combining a silicon MOSFET stack and a few layers of parallel nanowires connected by self-assembled molecular electronic devices. Such hybrids promise unparalleled performance for advanced information processing, but require special architectures to compensate for specific features of the molecular devices, including low voltage gain and possible high fraction of faulty components. Neuromorphic networks with their defect tolerance seem the most natural way to address these problems. Such circuits may be trained to perform advanced information processing including (at least) effective pattern recognition and classification. We are developing a family of distributed crossbar network (CrossNet) architectures that permit the combination of high connectivity neuromorphic circuits with high component density. Preliminary estimates show that this approach may eventually allow us to place a cortex-scale circuit with about \(10^{10}\) neurons and about \(10^{14}\) synapses on an approximately \(10 \times 10 \text{cm}^2\) silicon wafer. Such systems may provide an average cell-to-cell latency of about 20 nsec and, thus, perform information processing and system training (possibly including self-evolution after initial training) at a speed that is approximately six orders of magnitude higher than in its biological prototype and at acceptable power dissipation.

KEYWORDS: neuromorphic networks; VLSI circuits; nanowires; molecular self-assembly

INTRODUCTION

Ultradense integrated circuits with sub-10-nm features would provide enormous benefits for all information technologies, including computing, networking, and signal processing. However, recent research results indicate that the current VLSI paradigm based on CMOS technology and digital number crunching can be hardly extended into this region. Indeed, below 10 nm gate length the sensitivity of parameters (most importantly, the gate voltage threshold) of semiconductor field-effect transistors to inevitable size fluctuations grows exponentially. This sensitivity may send the fabrication facility costs (extremely high even now) skyrocketing and lead...
to the end of the exponential, *Moore’s Law*, progress\(^1\) some time during the next decade.

The main alternative nanodevice concept, single-electron devices\(^4\)-\(^6\) offers potential advantages over CMOS; in particular because (in contrast to field-effect transistors) their operation mechanism does not require high conductivity of device-to-electrode interfaces and, thus, allows a broader choice of possible device materials. However, the critical dimension of single-electron devices (the island size) for room-temperature operation should be below about 1nm\(^6\) far too small for current\(^1\) or even forthcoming\(^7\),\(^8\) lithographic techniques.

This impending crisis may be resolved by a radical paradigm shift from purely CMOS technology to hybrid *CMOL* circuits\(^3\). Such circuits (see Figure 1) would combine a level of advanced (e.g., 45-nm node\(^1\)) CMOS fabricated with the usual lithographic patterning, a few layers of parallel nanowire arrays formed (e.g., by nanoimprinting\(^9\)), and a level of molecular devices that would self-assemble on the wires from solution. A CMOL chip would combine the advantages of its nanoscale components (high reliability of CMOS circuits and miniscule footprint of molecular devices) as well as those of patterning techniques: the flexibility of the usual lithography and potentially very low cost of nanoimprinting and chemically directed self-assembly\(^1\)\(^0\)-\(^1\)\(^4\). This combination should enable unparalleled potential density CMOL circuits—up to \(3 \times 10^{12}\) active devices per square centimeter\(^3\). The recent spectacular demonstration of single-molecule single-electron transistors by several

![FIGURE 1. General concept of a CMOL (hybrid CMOS/nanowire/molecular) circuit.\(^3\)](image-url)
groups offers every hope that chemically-directed self-assembly of such devices with acceptable yield will be developed, hopefully in time to preempt the impending Moore’s Law crisis.

The most revolutionary part of CMOL are of course the single-molecule components. For these, single-electron devices are the leading candidates. For memory applications, two-terminal molecules with bistability due to conformation change is another possible option, but for any other electronics application the use of three-terminal devices is virtually unavoidable. The CMOS layer allows CMOL to circumvent one prominent drawback of single-electron transistors, their low voltage gain. However, the CMOL concept per se does not overcome another drawback of single-electron components (and virtually any nanometer-scale devices), namely, their high sensitivity to random charged impurities. Furthermore, chemically-directed molecular self-assembly hardly ever provides 100% yield of good devices. This means that the architecture of future CMOL circuits should be highly defect-tolerant. For circuits with a simple structure, such as memory matrices, such tolerance may be achieved in several ways, but for usual Boolean logic circuits no effective way around this effect has been suggested to our knowledge. For example, our estimates show that the Teramac-style approach leads to unfavorable performance scaling with the computing system size. On the other hand, an attempt to implement the redundancy in quantum-dot circuits (partly reminiscent of the CMOL concept) leads to a dramatic increase in the number of nanoelectronic components per function and the possible density advantage over nanoscale CMOS fades away.

Another motivation for a major revision of information processing techniques comes from the well-known comparison of the performance of present day computers and biological neural systems for one of the simplest advanced tasks—image classification. A mammalian brain recognizes a complex visual image with very high fidelity in approximately 100 milliseconds. Since the elementary process of neural cell-to-cell communication in the cerebral cortex takes approximately 10 milliseconds, this means that this task is completed in just a few “ticks” of the cortical circuitry. In contrast, the fastest modern microprocessor performing digital number crunching at a clock frequency of a few GHz and running the best commercially available software, would require minutes (i.e., on the order of $10^{11}$ clock periods) for an inferior classification of a similar image.

The contrast is very striking indeed, and calls for the development of biologically-inspired architectures for CMOL circuits. The requirements for these architectures are very stringent: on one hand, they should provide for advanced information processing (as a minimum, for image recognition and classification). On the other hand, the circuits should sustain the high physical characteristics of CMOL circuits, including density and speed, at acceptable power consumption. Several directions toward this goal have been explored (see, e.g., Refs. 24 and 27–31). Unfortunately, in our view, none of these approaches satisfies the requirements formulated (for details, see below). The goal of this paper is to describe our approach to this problem: the development of a family of neuromorphic architectures that we call distributed crossbar networks, or just CrossNets for short.
MOLECULAR DEVICES

The nanowire/molecular level of the CrossNet is a uniform field of latching switches (i.e., bistable devices) forming synapses between perpendicular axonic and dendritic nanowires. Currently, we are working on three-terminal switches that permit the implementation of Hebbian synapses. (Our initial attempts to use simpler, two-terminal latching switches ran into problems with their Hopfield-mode training. We maintain the hope that such devices may become useful again for the networks under continuous-mode training.) Figure 2 shows a possible single-electron implementation of such a switch. It consists of a single-electron transistor coupled electrostatically to another well-known device, a single-electron trap. This switch may be considered a three-terminal version of the four-terminal device discussed qualitatively in Reference 23. The transistor galvanically connects only two nanowires, one axon, and one dendrite. (Previous suggestions to use single-electron devices in neuromorphic networks were focused on the implementation of synapses merged with cell bodies—somas. However, we believe that the somatic

![Diagram](image)

**FIGURE 2.** Three-terminal single-electron latching switch: (A) schematics, (B) circuit notation (compare with Fig. 3), (C) result of simulation within the framework of the orthodox theory of single-electron tunneling, and (D) possible molecular implementation. C and R are the capacitance and resistance of each tunnel junction (shown as dark gray in A); Q are the background charges of the islands that may be fixed using additional global gates (not shown for clarity).
function has to be left for field-effect transistors, because high voltage gain, that is impossible in single-electron devices, seems necessary for the effective network functioning. The relatively low density of the somatic cells makes such implementations very feasible.) The effective voltage $V$ applied to the trap is contributed equally by two axons, one providing the input signal for the target cell $j'$, and another carrying the output signal of that cell,

$$V = \frac{V_j + V'_j}{2} = \frac{V_0}{2} \times (\pm v_j \pm v'_{j}),$$  \hspace{1cm} (1)$$

where $v_j$ and $v'_{j}$ are normalized output (axonic) signals of the cells with values in the interval $(-1, +1)$, and the signs of operands in the parentheses depend on the axon polarity (see CrossNets below).

Figure 2C shows the results obtained from Monte Carlo modeling of the device properties, within the simplest orthodox theory of single-electron tunneling. This theory is quantitatively valid for relatively large single-electron islands, with quasi-continuous electron spectrum. (For molecular devices, where the effective size of the island is very small (less than 1 nm) and the electron energy quantization is essential, the theory should be modified. However, this modification does not change the results qualitatively; for example, the observed $I$–$V$ curves of molecular single-electron transistors are close to those following from orthodox theory.)

If voltage $V$ is low, the trap in equilibrium has no extra electrons and its total electric charge $Q = -ne$ is zero. As a result, the transistor remains in the Coulomb blockade state, and input and output wires are essentially disconnected—see the horizontal line in Figure 2C. If $V$ is increased beyond a certain threshold $V_{inj}$ (which is lower than the Coulomb blockade threshold voltage $V_t$ of the transistor), one electron is injected into the trap: $n \to 1$. In this charge state the Coulomb blockade in the transistor is lifted, connecting the wires with a finite resistance (for a symmetric transistor, close to the tunnel resistance $R$ of a single tunnel barrier) at any $V$. Only when the cell activity (voltage $V$) becomes low for a long time do, eventually, either the high-order quantum processes (cotunneling) or unavoidable thermal fluctuations kick the electron out of the trap, closing the transistor and disconnecting the wires.

Figure 2D shows a possible single-molecular implementation of the latching switch. Arenediimide groups were chosen for single-electron islands due to their strong electron acceptor properties. (They have already been exploited in the design of a variety of molecular materials.) In particular, naphthalenediimide, the acceptor group used in our design, forms a stable radical anion. The transistor island is connected to two electrodes by means of short conducting oligo-phenylethylenylene (OPE) chains that are terminated by thiolate alligator clips that allow self-assembly of the molecule on the nanowire electrodes. A parallel conducting chain connects the single-electron trap and the axon running from cell $j'$ is an insulating structural element; it keeps the trap at a fixed distance from the wire from cell $j'$ and stabilizes the geometric arrangement of the synapse. The benzyl-aryl ether links between the two conducting wires are also insulating structural elements; they control the distance between the trap islands and the transistor island. In the proposed configuration of the switch, the acceptor group 3 used for the terminal island of the trap is effectively stronger than the group 1 of the transistor, as a result of being electronically coupled to the pyromellitimide group 2.
It is important that the synthesis of such molecules may be executed entirely by means of well established chemical transformations. The diimides are formed from the corresponding anhydrides and aromatic amines. The OPE chains are generated via the palladium-catalyzed coupling of aryl halide and arylethynyl components and are connected to the diimides by means of the same type of coupling process. The benzyl-arylether links between the conducting chains are established at appropriate stages of the construction process by condensation of phenol and benzyl bromide intermediates. A bicyclo[2.2.2]octane group is inserted in the link between the islands of the trap and the wire from cell 1 in order to make this segment electronically insulating. Arenediimides, as well as extended OPE chains, possess low solubility in most common solvents. For this reason, alkyl substituents, for example, hexyl or isopropyl groups, are introduced on most of the arene rings to make the products tractable. (These substituents are omitted from FIGURE 3D for clarity.) The molecular design fits the nanowire electrode geometry, enabling chemically-directed self-assembly of the molecules from solution. Our group has already synthesized the single-electron transistor part of this molecule, and we are working on their self-assembly and transport characterization.

In future, the switch design may be further improved. For example, island 2 of the trap may be replaced by a thicker tunnel junction (e.g., a longer OPE chain), making electron escape time from island 1 sufficiently long.

**CROSSNETS**

FIGURE 3 shows the general idea of the CrossNet using as examples its two species, inclined distributed crossbar (InBar, Fig. 3A) and randomized distributed crossbar (RandBar, Fig. 3B). Figure 3C shows the structure of an elementary cell of this field, the synaptic tile (or plaquette). The structure may seem complicated, but in fact is straightforward, essentially the simplest result of symmetrization and localization of the well-known crossbar (or crosspoint) switching networks: each plaquette is crossed by just two axonic and two dendritic wires, of opposite signal polarities, going in each of four directions available on a square lattice. Physically, the axonic and dendritic wires may be similar; they differ only in how they are connected to the cell bodies (somas).

The somas (FIG. 3D) may be implemented as MOSFET voltage amplifiers with a sigmoid-type saturation (activation) function $v_j = g(u_j)$ common for the fire-rate models of neural networks. (Here $u_j = \sum w_{ij}' v_{j'}$ is the input signal of the amplifier, physically the voltage difference between two incoming dendrite wires.) An important role is played by the passive (open-circuit) terminations of incoming axonic and outcoming dendritic wires at the somatic cell interfaces. These terminations do not allow direct interaction of two somatic cells located in the same row or column of the CrossBar and limit the length of dendritic and axonic wires and, hence, the cell connectivity parameter $M$. (This number is defined as a ratio of the number of synaptic plaquettes to the number of somatic cells in the array.) As a result, in InBar, each cell is connected to exactly $4M = 4\tan^2 \alpha$ near neighbors (they are all located inside a square inclined at angle $\alpha$ relative to the synaptic plaquette array). In RandBar, with its random location of somatic cells, the lengths of axonic
and dendritic branches obey the Poisson statistics with mean $M$, and each cell is connected to an average of $4M$ other cells. Thus, the connectivity of all CrossNets is quasilocal and at $M \neq 1$ these networks occupy the design niche in-between cellular automata (with only adjacent cells directly connected), and fully connected Hopfield-type networks.\(^5\)

**FIGURE 3.** (A) InBar-type and (B) RandBar CrossNets, and schematics of their (C) synaptic plaquette (white cell) and (D) somatic (gray) cell. *Solid lines* show axonic and *dashed lines* show dendritic nanowires. *Dots with arrows* are three-terminal switches (Fig. 2). Signs on the top of (C) show the output polarity (for axonic wires) and destination input polarity (for dendritic wires). Due to the rotational symmetry of the synaptic plaquette, these signs are shown for one side only. *Solid points* in (D) show open-circuit terminations of the nanowires. Due to these terminations, somas of the same row or column do not interact directly. If the load resistors $R_L$ are sufficiently low (much less than $R/M$), CrossNets tolerate synapse self-assembly at all nanowire crossings.
A closer examination of Figure 2 shows that each two somas are directly connected by (maximum) four four-switch groups (see Figure 4). The main goal of this grouping is to convert the linear signal addition expressed by Eq. (1) into the signal multiplication using the synapse nonlinearity. We explain this point in more detail. The connection in single-electron latching switches is not deterministic but rather probabilistic. (This is the main feature of all single-electron devices.3–6) If voltage $V$ does not approach $V_{inj}$ too closely, the probability $p$ of the trap being filled (and hence, the synapse connected) may be described by the master equation

$$\frac{dp}{dt} = \Gamma_\uparrow(1 - p) - \Gamma_\downarrow p,$$

where, with a good precision, the switching rates may be described by the Arrhenius law

$$\Gamma_\uparrow = \Gamma_0 \exp\left\{\frac{e(\pm V - S)}{k_BT}\right\}.$$  

Here, $T$ is the effective temperature and $S$ is a shift that may be changed by a voltage applied to the trap island via an additional global (or quasi-global) gate. If not addressed properly, switching randomness leads to additional uncertainty (digital

![Figure 4](image_url)

**FIGURE 4.** Direct connection of two cells in a CrossNet array. (Only one way connections are shown for clarity; feedback connections are similar.) The axons shown by short lines come from the corresponding outputs of cell $j'$. Each group of four latching switches may be used as a Hebbian synapse.
noise) that may have a detrimental effect on information processing.\textsuperscript{55–60} However, the strong nonlinearity of Eq. (3) makes it possible to suppress this randomness almost completely (quasifuzzy connections). We demonstrate how this may be achieved using as an example the Hebb rule implementation.

The classical Hebb rule\textsuperscript{55–59} (obeyed by most synaptic connections in the cerebral cortex, see, e.g., Refs. 25 and 59) requires synaptic weight to increase when the presynaptic (axonic) and postsynaptic (dendritic) activities coincide in time. In software-implemented neural networks this is achieved by making the synaptic weight $w_{jj'}$ (or its time increment) proportional to $v_j \times v_{j'}$. The connectivity of a single switch depends not on the product, but on either the sum or the difference of $v_j$ and $v_{j'}$—see Eq. (1). However, let the signals be strong ($v_j \approx \pm 1$) and constant in time. Then the solution of Eqs. (2) and (3) at $t >> 1/\Gamma_0$ is

$$p = \frac{1}{2} \left[ 1 + \tanh \frac{e(V - S)}{k_B T} \right].$$

(4)

If shift $S$ satisfies the condition $k_B T << eS << eV_0$, then only one of four synapses (that having $\pm v_j \pm v_{j'} = \pm 2$) will have connection probability $p$ very close to 1, the other three will be almost certainly disconnected ($p = 0$). Hence, the signal current sent to the target cell’s dendrite is virtually deterministic and very close to ($V_0/R) \times \text{sgn}(v_j \times v_{j'})$. This property may be interpreted as synaptic weight $w_{jj'} = \text{sgn}(v_j \times v_{j'})$, corresponding to the discrete (clipped\textsuperscript{55}) Hebb rule.

Thus, the four-switch groups may be used as binary-weight, analog-signal Hebbian synapses. The binary character of the synaptic weight does not impose too heavily on the neural network performance.\textsuperscript{55–60} For example, the storage capacity of the Hopfield network with binary (clipped) synaptic weights is only about 30\% lower than that with continuous weights.\textsuperscript{55,56} For single-layer perceptrons the weight clipping-induced difference is also small (logarithmic in the processing error).\textsuperscript{50}

**HOPFIELD-MODE TRAINING**

We have used this property to demonstrate\textsuperscript{34,35} successful training of InBar as a Hopfield (we use this commonly accepted name despite the considerable controversy concerning the authorship of this concept\textsuperscript{61}) network.\textsuperscript{34} During this demonstration, we self-imposed all the restrictions anticipated for the future hardware CrossNet implementations. Most importantly, no external access to an individual synapse is possible, other than via the corresponding somatic cells. (In InBar, access to each cell is especially convenient, since all somas may sit in nodes of a rectangular crossbar matrix—see Fig. 3A.) This limitation, as well as the deeply recurrent structure of CrossNets, does not allow the straightforward use of most previously developed techniques of neural network training.\textsuperscript{55–59}

We developed two techniques that allowed us to overcome these limitations. The first (pixelization) worked reasonably well for InBar with a large connectivity parameter $M$, but the maximum number $P$ of patterns we could record with good restoration fidelity was rather small. An analytical estimate has shown that $P_{\text{max}}$ (known as network capacity) should in fact be small, on the order of $M^{1/2}$. (This estimate should be compared with the well-known result $P_{\text{max}} \sim 0.1 N$ for the fully connected Hopfield network of $N$ cells.)
However, the second technique (external multiplication) gave a much better result. In this method each pair \( \{j, j'\} \) of cells is taught in turn. (Due to the CrossNet quasilocality, cell pairs separated by Manhattan distance \( L_{jj'} > M \) may be trained simultaneously.) For this, one cell of the pair is fed with the strong external signal \( V_{0p} \sum_{p} \xi_{j}^{p} \xi_{j'}^{p} \), where \( \xi_{j}^{p} \) is the \( j \)th pixel of the \( p \)th pattern of the training set, and the second cell is fed with positive signal \( V_{0} \). In this way, each of two four-switch synapses connecting the cell pair is exposed to training only once and, as described in the end of the CROSSNETS section, probabilities of connection of its synapses are saturated to provide virtually deterministic weights \( w_{jj'} = sgn \sum_{p} \xi_{j}^{p} \xi_{j'}^{p} \). This is the so-called “clipped Hebb rule” that is known to work very well for fully connected Hopfield networks.\(^{55,56}\) Our estimates\(^{35}\) show that for a CrossNet trained by this method, \( P_{\text{max}} \) should scale as \( M \); this is as much as could be expected from the well-known theory of randomly diluted networks.\(^{55,56}\) (The locality of CrossNets is to some extent similar to dilution.)

Numerical experiments with InBar computer models have confirmed this result. An example of our results is shown in FIGURE 5. In this case, an InBar had been taught three different images and was then able to restore a strongly corrupted version of any of these images fed into it as an initial condition. Remarkably, the recall process is perfect (no corrupted bits in the end) and very fast. It is important to emphasize that such performance requires a large connectivity parameter \( M \neq 1 \) and is impossible in cellular-automata-type neural networks\(^{24,30,62,63}\) with nearest neighbor interactions and, hence, low connectivity.

**PROSPECTS FOR CLASSIFICATION TRAINING**

We are of course happy with these initial results, but understand that much more has to be done, because practical applications of Hopfield networks are rather limited (essentially to associative memories). Much more important is another type of information processing, pattern classification.\(^{57–59}\)
The main problem we face in this direction, besides inaccessibility of individual synapses, is the deep recurrence of CrossNets. Generally, this property can lead to system latch-up in one of its spin-glass states, making the effective processing of information impossible. Fortunately, in CrossNets the latch-up problem may be overcome because of the sign symmetry of its synaptic weights (Fig. 4). In fact, let the synaptic connections be completely random, with the same connection probability \( p \) for all synapses. (This is easy to achieve, for example, by making the global shift \( S \) small and making the somatic amplifier saturation level relatively low, \( eV_0 = k_B T \), in this case all \( p \) will rapidly drift to the same level, 0.5). In this case the ensemble average \( \langle w_{ij} \rangle \) of the synaptic weights is zero. This means that for any closed signal loop passing through several neurons there is an equal chance to have a total positive feedback and total negative feedback (at zero frequency). Loops with positive feedback, of course, tend to form local dc latch-ups if the effective somatic cell gain \( g \equiv G_{RL}/R \) is high enough. However, at a finite frequency \( \omega \), each intercell connection introduces an additional phase shift \( \Delta \phi = \arctan(\omega \tau_0) \). Therefore, at sufficient gain, self-excitation takes place in the negative-feedback loops as well, but at a finite frequency \( \omega = 1/\tau_0 \). At high connectivity, \( M \neq 1 \), dc latch-ups and ac oscillations in various loops interfere in the same cells and, due to the somatic activation function nonlinearity, interact with each other. At relatively large \( g \) this interaction should lead to chaotic dynamics.

Our analysis and numerical experiments\textsuperscript{32–34} confirm these expectations (see, e.g., Figure 6 and Figure 7). Notice that near the critical value \( (g_c = M^{-1/2}) \) of the gain, local dc latch-ups develop first. On further increase of gain, the system develops quasisinusoidal ac oscillations. At larger gain, more and more oscillations are developed and at \( g = 1.5g_c \) the process is almost completely chaotic.\textsuperscript{64} Figure 7

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure6.png}
\caption{CrossNet self-excitation: typical time dependences of normalized signals \( x_i \) of two randomly selected dendrites of a 2,500-neuron RandBar with \( M = 16 \) and \( p = 0.5 \), for several values of the ratio of the somatic gain \( g \) to its critical value \( g_c \).}
\end{figure}
shows that the special distribution of neural activity at such self-excitation is also virtually random. At larger $g$, the chaos is apparently robust.$^{65}$

In the light of this property, a possible method of continuous network training may be as follows (this is essentially a global twist of the known reinforcement learning techniques—see, e.g., Ch. 9 of Ref. 59). Suppose that, in a CrossNet with $N \neq M$ neurons we designate a relatively small number $O \approx N/M$ of neurons as output cells and a larger number $I (O \ll I \ll N)$ neurons as input cells, leaving all other cells to play the role of a huge hidden “layer”$^{57-59}$ (although this term is hardly applicable to deeply recurrent networks, such as CrossNets). Setting the initial shift $S$ to zero, we ensure that all synapses are random, so that by increasing $g$ above $g_c$, we get into the chaotic regime. Now impose on the input cells the first image of the training set. As soon as the output signals match, or correlate considerably with our desired output, we increase the (quasi-)global shift $S$ to about $V_0$ for a short time $\Delta t = 1/T_0$, allowing a partial synaptic adaptation. By repeating this procedure for all patterns, we should arrive at the system with fixed synaptic weights (where the chaos disappears) that hopefully will provide effective image classification.
There are still many open questions about the viability of this idea. For example, since the available phase space is larger than the phase space of the desired input/output combinations, it might take an impractically long time for the system to learn. This is why the analytical study and numerical-model testing of this idea is the first priority of our CrossNet architecture work. The next important task would be the defect tolerance characterization of this operation mode. We are also planning to explore an alternative way of training using a dual CrossNet for error backpropagation.

**DISCUSSION**

A possible success of the chemically-directed assembly of molecular synapses and CrossNet model training would justify a large-scale industrial effort toward practical implementation of the first CrossNet circuits. The main feature of these circuits, initially of relatively small integration scale (about $10^6$ to $10^7$ neurons), would be their unparalleled speed performance. In fact, let us estimate the speed assuming the nanowire spacing (half-pitch) of $F = 2$ nm, limited by wire-to-wire tunneling. For a CrossNet with connectivity $4M = 10^4$, the Hebbian synaptic plaquette size would be $32 \times 32$ nm$^2$, corresponding to the dendrite wire capacitance $C_0 = 3aF$ per plaquette. Note that this corresponds to an areal density of about $10^8$ cells per cm$^2$ (at $n = 1.5 \times 10^{12}$ synapses per cm$^2$), higher than that of the cerebral cortex, despite the quasi-three-dimensional character of the cerebral cortex versus the quasi-two-dimensional structure of CMOL circuits (Fig. 1). For the molecular implementation of the synapses, the time scale $\tau_0$ of interaction of two neural cells is dominated by charging these capacitances through relatively high resistances $R$ of single-electron transistors: $\tau_0 = RC_0/4$. The reduction of $R$, and hence $\tau_0$, is limited by an acceptable power dissipation per unit area that is close to $nV_0^2/R$. For room-temperature operation, the voltage scale $V_0$ should be of the order of 1 volt to avoid thermally-induced errors. With the high but acceptable power consumption of 100 W/cm$^2$, achieved at $R = 10^{10}$ $\Omega$ (a realistic value for the devices like that shown in Fig. 2), we get $\tau_0$ as small as 20 nsec. This speed is approximately six orders of magnitude higher than that of the cerebral cortex circuitry. Even scaling $R$ up by a factor of 100 to bring power consumption to a more comfortable level of 1 W/cm$^2$, would still leave us with four orders of magnitude of speed advantage. This is why we believe that even relatively small CrossNet chips may revolutionize the image classification field, including such important applications as handwriting and speech recognition and detection of optical image features.

This success would pave the way toward much more ambitious goals. It seems completely plausible that a cerebral-cortex-scale CrossNet-based system (with about $10^{10}$ neurons and $10^{14}$ synapses, that would require about $10 \times 10$ cm$^2$ silicon substrate) would, after a period of initial training by a dedicated external tutor (see Fig. 8B), be able to learn directly from its interaction with the environment. Such large-scale systems would require a hierarchical organization involving, at least, the means of fast signal transfer over long distances (Fig. 8A). Fortunately, for the InBar-type CrossNet with its regular location of somatic cell interfaces, such communication is easy to organize (Fig. 8B). In this case one can speak of a “self-evolving” system. This idea is philosophically close to, but in practice rather different from the
FIGURE 8. A. Possible hierarchical organization of a large-scale, CrossBar-based information system. B. A simple scheme for fast (though relatively rare) communication of distant neural cells in that system.
earlier concepts of evolutionary computation,67 evolutionary algorithms,68 evolving hardware,69 and global evolving artificial neural networks.70 (The main distinction is, again, the impossibility of direct access to an arbitrary synaptic weight in CMOL-based CrossNets.) If these expectations are confirmed, we may be able to revisit the initial dream of neural network science of providing hardware means for reproducing the natural evolution of the neocortex on a much faster time scale. Such evolution may lead to self-development of such advanced features as system self-awareness (consciousness) and reasoning. If a substantial success along these lines materializes, it will have an enormous impact, not only on information technology, but also on society as a whole.

ACKNOWLEDGMENTS

Fruitful discussions with P. Adams, J. Barhen, S. Grossberg, V. Protopopescu, and T. Sejnowski are gratefully acknowledged. The work was supported in part by ARDA via ONR, DOE (directly and via ORNL), and NSF. Most numerical calculations were carried out on the Njal supercomputer cluster that had been acquired with a grant from DoD’s DURIP program via AFOSR.

REFERENCES


