*J.I. Raffel, J.R. Mann, R. Berger, A.M. Soares, and S. Gilbert*

# A Generic Architecture for Wafer-Scale Neuromorphic Systems

Neural-network architectures emulate the brain by using a parallel, highly interconnected system of simple processing units to solve problems. Systems based on neural networks offer promising solutions to complex problems in recognition, control, and robotics. However, the massive parallelism and high fan-out of neural networks impose enormous interconnection requirements on the integrated-circuit designs that implement them. Wafer-scale integration interconnects many circuits on a single wafer, thus eliminating wirebonds, package pins, and external printed-circuit wiring. A generic wafer-scale device for neural networks, which uses multiplying digital-to-analog converters for programmable synapses and operational amplifiers for summing nodes, has been developed. After each wafer is fabricated, laser cuts and links may be used to define the network connectivity and provide defect avoidance for yield improvement.

Wafer-scale integration is ideally suited to handling the interconnection requirements of neuromorphic architectures. The pin-outs of conventional integrated-circuit packages do not provide enough fan-out capability to permit the implementation of these highly connectionist architectures. But, since wafer-scale integrated circuits interconnect several circuits on one wafer, large numbers of package pins are not necessary.

Wafer-scale integration solves the pin-out issue, but it also creates two problems: wafer defects and fabrication costs. Silicon wafers cannot be made defect-free across large areas, so defects must be circumvented. And because wafer-scale systems are so complex, fabrication of small quantities of custom devices can be extremely expensive. (For a discussion of these issues, please see the box, "Wafer-Scale Integration.")

The restructurable very large-scale integration (RVLSI) approach to wafer-scale integration solves both the defect problem and the fabrication problem [1]. In the RVLSI approach, one generic design is used for the fabrication of all ICs. Then, after the generic ICs have been fabricated and tested, lasers are used to customize devices — by forming and removing conductors.

RVLSI solves the defect problem because it permits designers to place redundant cells on wafers and then to isolate defective cells. And, because only one generic set of photolithographic masks is required, fabrication expenses are reduced.

A generic wafer-scale architecture, which will be capable of implementing a wide variety of neuromorphic designs, such as Hopfield nets and multilevel perceptrons [2], is being developed at Lincoln Laboratory. (An introduction to these systems is included in the box, "Neural Networks.")

## Restructurable Very Large-Scale Integration

The critical technological advance that underlies the RVLSI approach to building wafer-scale circuits is the development of a very high-yield laser-linking technique (Fig. 1). The laser-linking techniques permit both additive and deletive modification of interconnect wiring after wafers have been fabricated and tested. RVLSI allows the incorporation of redundant devices, and the removal of defective ones, without the use of time-consuming and defect-prone custom masks or extra fabrication steps.

A number of link structures have been developed; two have been demonstrated in whole-

# Wafer-Scale Integration

Wafer-scale design begins with the definition of a building block small enough to provide high yields, but large enough to keep the overhead area required for restructuring at a tolerable level. Once the building block is defined, it is placed in an array within a matrix of conductors and programmable links.

Figure A shows a schematic representation of an array of such building blocks. In this array, $m$ input lines enter each block horizontally and $n$ output lines exit each block vertically. Unlike a discrete-package implementation, the values of $m$ and $n$ are not limited by the pin-out; they are determined by the fan-in and fan-out requirements of the circuit. To provide complete flexibility in the configuration of connections between the outputs of any block and the inputs of any other blocks, a matrix of links and conductors of size $j \times n$ is located below each block.

Because signals are not routed off the wafer (and across a printed-circuit board), lead inductances, stray capacitances, and printed-circuit–board noise are avoided. Therefore, wafer-scale devices are intrinsically faster and less sensitive to noise.

The major impediment to wafer-scale integration has been the lack of a scheme that deals with the defects occurring on all integrated-circuit wafers. Restructurable very large-scale integration (RVLSI) solves this problem — it uses laser linking and cutting to connect the working segments of a wafer and to isolate the defective segments. In the RVLSI technique, custom devices are constructed on wafers only after a generic wafer-scale circuit has been fabricated and tested. Therefore, RVLSI maximizes the usage of each wafer.

RVLSI technology has been successfully used at Lincoln Laboratory. Devices fabricated with the RVLSI method include a 16-point FFT processor [4], a Hough Transform Processor, and a device that implements the Dynamic Time Warping [5] algorithm for a speech recognizer. The repetitive nature of signal-processing applications translates into a regular, repetitive architecture for the signal-processing device, and is thus ideal for wafer-scale implementation.

Like signal-processing architectures, neural-network (or connectionist) architectures feature regular, repetitive operators arranged in arrays that require a high degree of connectivity. Current connectionist architectures are usually considered highly redundant, and therefore resistant to individual component failures.

But the systems that are likely to be built in the near future will have far less redundancy than their neurological counterparts and will thus require the capability of testing interconnections and active devices completely.

In the wafer-scale neural-network devices developed at Lincoln Laboratory, each weighted internode connections is a multiplying digital-to-analog converter and each node is a summing amplifier. To provide good testability, each MDAC is accessible from the wafer periphery (Fig. A). Furthermore, blocks are extendable, so fan-in and fan-out can be increased in increments of the block width and length.

Horizontal input lines and vertical output lines are continuous across block boundaries, and extend completely across each wafer. The input and output lines therefore give access to the blocks and allow complete testing of interconnections.

Above each $j \times n$ link matrix is a row of amplifiers. The outputs of the amplifiers are connected to the common horizontal line, "Amplifier Test Out." The inputs of the amplifiers are connected to the vertical line, "Amplifier Test In," which is shared by all amplifiers in a column. After testing, the common connections between amplifiers can be cut away.

wafer systems. The first structure, a vertical device (Fig. 2), comprises two levels of metal separated by a dielectric sandwich of amorphous silicon covered on top and bottom by oxide barriers. The oxides prevent interdiffusion of metal and silicon during high-temperature annealing steps. The resistance of the device in Fig. 2 is under 1 $\Omega$ after programming. Device yields are extremely high over a wide range of laser power. Whole-wafer systems developed at Lincoln Laboratory with the use of the vertical-link technology include a 16-point FFT [2], a Hough Transform Processor, and a two-dimensional convolver. The drawback of the vertical device is that it requires unconventional processing steps, and thus is not easily processed in standard industrial lines.
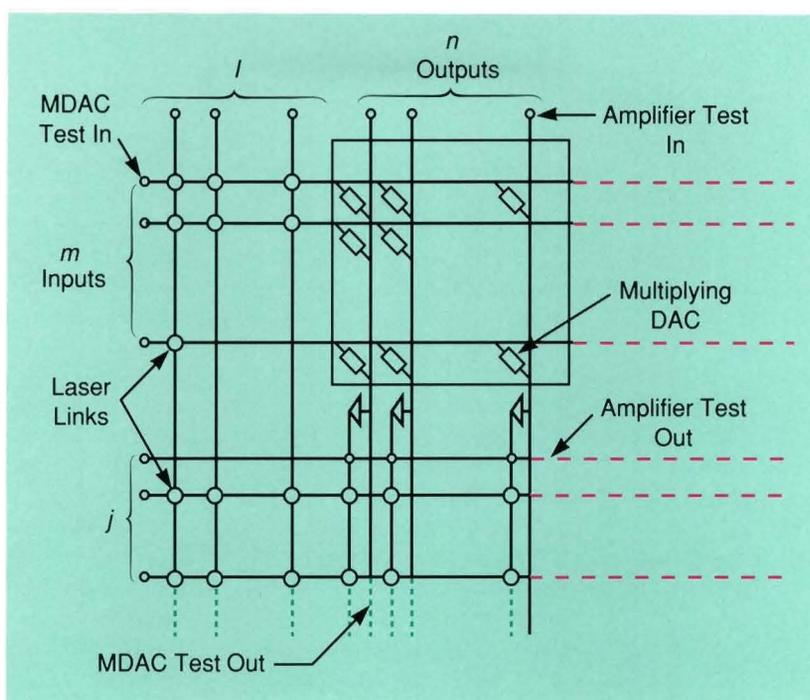
A second link structure has been developed,

*Fig. A — A flexible neural-network application can be implemented with this restructurable VLSI array of mutiplying digital-to-analog converters (MDAC), summing amplifiers, and laser links.*

and this structure is suitable for standard semiconductor processing lines. The new link requires no additional processing steps and can be fabricated in a conventional CMOS process sequence. This device, shown in Fig. 3, is a lateral structure. It is formed by two substrate diodes separated by a gap approximately the size of the laser beam (typically 2 to 4 $\mu$m). Exposure at power levels and pulse widths comparable to those used for the vertical link (Fig. 2) causes the silicon in the gap to melt and the dopant to diffuse across the gap, resulting in links in the 100-$\Omega$ range. The laser system can also be used to segment wires on both first- and second-level metal.

Recently, the lateral link was demonstrated in a Dynamic Time Warping wafer for speech recognition. This device was fabricated entirely in the MOS Implementation System (MOSIS) foundry, a dramatic demonstration of the ease with which this technology may be transferred to a standard semiconductor processing line.

Depending on the details of a neural-net application, either the vertical or the lateral links could be used to design a wafer-scale circuit. Since neural-network systems typically operate at high impedance levels and at low currents, the higher resistance of the lateral links will not pose a significant drawback.

To automate the linking procedure, a set of special-purpose CAD tools has been developed. By using the tools illustrated in Fig. 1 and the software that we have developed, most neural-net architectures can be implemented with only a minimal amount of additional software.

## Chip-Level Design

As a first step in evaluating a wafer-scale neural-network design, a discrete integrated circuit, which consisted of 32 analog input lines connected through an array of 4-bit multiplying digital-to-analog converters (MDACs) to 16 pairs of output lines, was designed, fabricated, and

# Neural Networks

Neural networks are computing elements that take a cue from biological neural systems and use massively interconnected arrays of relatively simple processors. Typical applications for neural networks include correlation, pattern recognition, and control. Most current work on neural networks is performed by using computer simulations of systems, but the confluence of revolutionary advances in integrated-circuit technology and growing interest in neuromorphic systems is now leading to electronic circuits that implement these systems.

The most commonly modeled neural nets are based on a model of a neuron, as shown in Fig. A, that produces an output dependent on the weighted sum of inputs from its peers. A wide variety of architectures, for such applications as character and speech recognition, have been built around this basic building block. The Gaussian classifier of Fig. 8 is one example of these architectures. An excellent summary of these architectures is given by Lippmann [2], and a textbook by Rumelhart and McClelland [6] is devoted to the subject.

Virtually all neural-network models include weighted connections; the building blocks that we've chosen use multiplying digital-to-analog converters (MDAC) to provide the weighting. The output current of each MDAC is the product of the input voltage and the total conductance of the converting transistors that are on. The number of transistors turned on is controlled by the stored bits, which represent the weight or strength assigned to a particular connection [7]. The output current of each of the MDACs is routed to a transconductance amplifier, which serves as the summing node found in neural-network models.



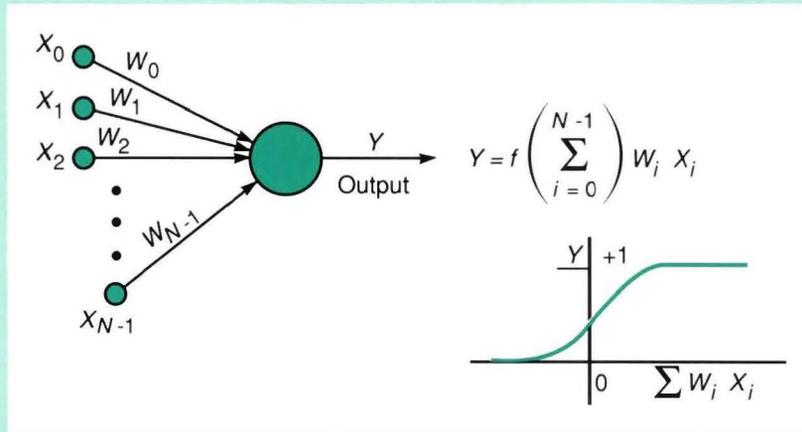$$Y = f\left(\sum_{i=0}^{N-1}\right) W_i \, X_i$$

*Fig. A — This basic model of a neural network has been used to simulate or develop nearly all the current neural-network models. The neuron model is shown here with a sigmoidal output, but a variety of outputs are possible. The input values are labeled $X_0$ through $X_{N-1}$, the connection weights are $W_0$ through $W_{N-1}$, and the output of the network is Y.*

tested (Fig. 4). The MDACs provide the weighting; off-chip transimpedance amplifiers convert the summed current to a voltage that neural network inputs require. Since the laser links had already been well characterized, the objective of this design was to test the active circuits — the MDACs — not the laser links that would be used in the wafer-scale design.

Figure 5 is a circuit diagram of an MDAC pair. One MDAC contributes current to an excitation output, $I_E$; the other contributes to an inhibition output, $I_I$. The $E/I$ control bit determines which of these is active. The weight transistors that perform the digital-to-analog conversion provide conductances that vary by factors of two ($1/2\ g_{ds}$, $1\ g_{ds}$, $2\ g_{ds}$, $4\ g_{ds}$). The configuration of the transistors determines the conductances, with two transistors in series for the least significant bit and up to four in parallel for the most significant bit.

Figure 6 shows a plot of output current vs. input voltage for a single MDAC. Each of the 15 traces gives the increase in current output associated with a stored digital value as the input voltage increases from 0 to 1 V. Measurements show that the output-current increments for all values of $V_{in}$ and all 16 values of the stored digital word are uniform to within 5%.

By tying input lines in parallel, the precision of the MDACs can be extended beyond four bits. For example, when two input lines are tied together, and all four bits in one of the MDACs are held at zero while the other MDAC's four bits are increased from zero to 15, the output current is increased from $I_0$ to 15 $I_0$. Once the output current has reached 15 $I_0$, the other MDAC can be incremented to provide an additional 15 $I_0$, doubling the output current in 1-bit steps and, therefore, providing an extra bit of precision. And the same principle can be used to increase the precision of the MDACs further — by doubling the number of parallel inputs for each additional bit of precision desired. One of the analog inputs can be connected to a fixed volt-
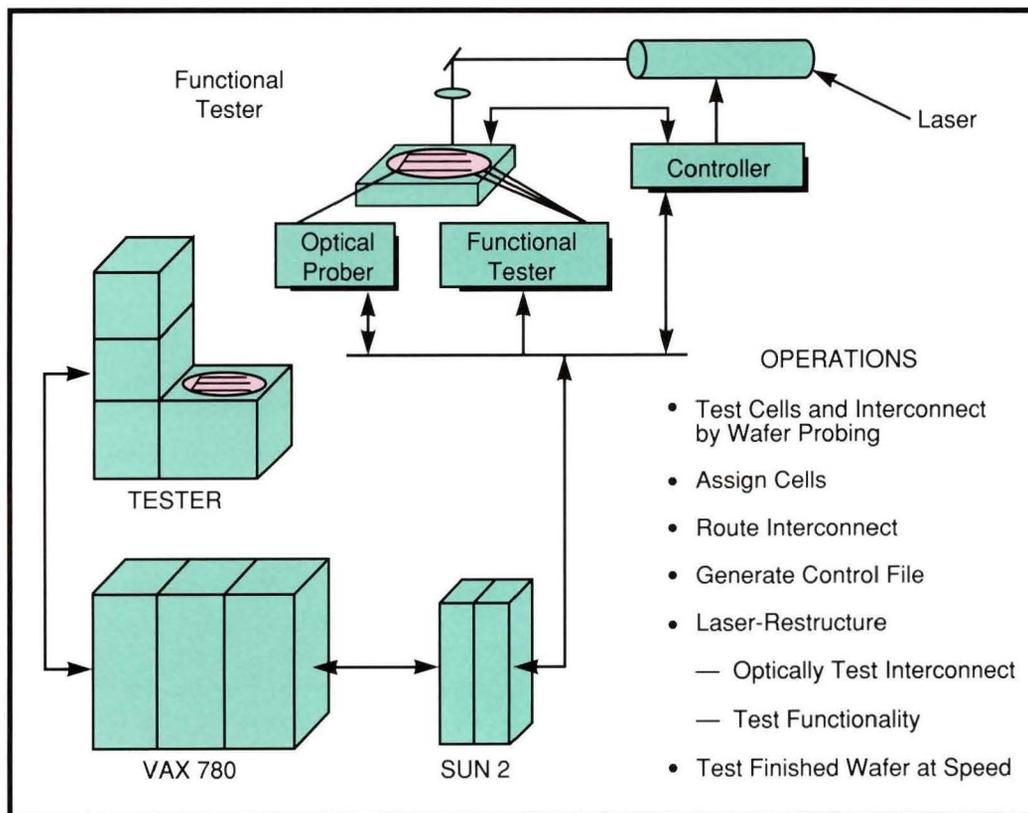


Fig. 1 — Laser facility used in the production of RVLSI circuits. The 2-W beam from the CW argon laser is focused on a silicon substrate through a rotating shutter that produces 1-ms pulses. The tested wafer is moved into position under the laser and circuit modifications are made at a rate of 6 additions/ deletions per second.
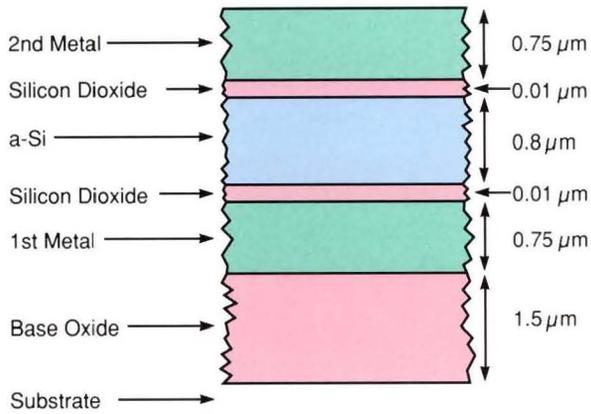
*Fig. 2 — The vertical link provides a low-resistance (<1 Ω) connection for laser-programmed circuitry.*

age, thereby providing threshold offset control for the off-chip amplifier that connects to the synapse outputs.

## Experimental Results

The weighted-summing circuits were characterized by building a Hamming net, which used two of the MDAC chips. The Hamming net, an optimum classifier for binary patterns corrupted by noise, enabled us to measure the uniformity of the weights and to verify the operation of a winner-take-all, or max-picker, circuit (i.e., a circuit that chooses the output with the highest value). The net was broken into two stages, a linear correlator and a max-picker circuit.
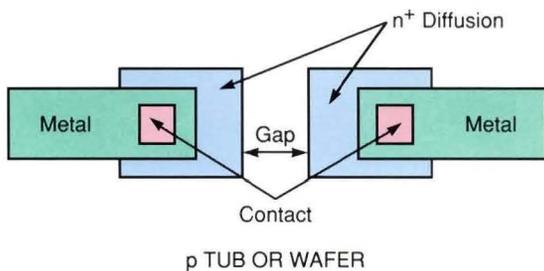


*Fig. 3 — Although the lateral interconnection circuit has a relatively high resistance associated with it (≈100 Ω), it requires no nonstandard fabrication steps and can be produced with conventional CMOS technology.*

The linear correlator acts much like an associative memory; when a vector is presented to its inputs, the correlator determines how well the input vector matches vectors stored within the synapse weights. The circuit has 16 input nodes, each of which receives a single bit of the input vector, and six output nodes, each with weighted connections to all input nodes. Weights are coded as +1 (excitatory) for a stored One and −1 (inhibitory) for a stored Zero. By applying a bias of −1 for each stored Zero at each node, the output can be made to produce a voltage that decreases from its maximum by an amount proportional to the Hamming distance. The maximum output, which is produced by the best match, therefore obeys the equation

$$n\text{th output} = k\,(16 - \text{Hamming distance to } n\text{th stored vector}).$$

The four-input two-output array shown in Fig. 7 illustrates the coding scheme. Two stored values, 1010 and 1110, are shown for the two output nodes. Their corresponding weight representations and the sum of products are generated by a perfect match with the input vectors. For the experimental system with 16 inputs and six outputs that is a direct extension of the $4 \times 2$ array, nine vectors were presented, of which six 16-bit vectors were identical to the six stored values. Three others had various degrees of mismatch.

The results of the experiment are shown in Table 1. The resulting outputs follow the equation above, with a variation of less than 5% for a value of $k$ equal to approximately 70 mV. Each of the six output nodes (labeled $A$ through $F$ in Table 1) was associated with a stored vector of the same label. When a vector similar to the stored vector was presented to the correlator, the corresponding output neuron (transimpedance amplifier) had, as planned, the highest output voltage. For example, when input vector 1, which was most similar to the A stored vector, was presented to the network, the output voltage at node A was 1,130 mV; the other nodes had significantly lower outputs.

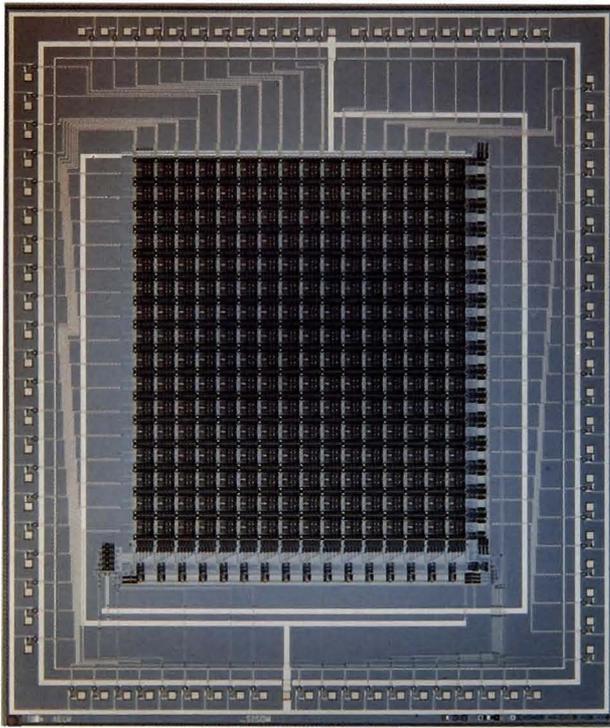The circled numbers along the matrix's diago-

*Fig. 4 — This chip was used to test the MDAC-based neural-network concept. The chip is packaged in an 84-pin package and uses 64 pins (32 plus 2 × 16) for analog inputs and outputs, five pins for digital data in, and 10 bits for MDAC address. The device was built with 3-µm design rules and occupies an active area of 28 mm $^2$. At present, 28,500 transistors are required to implement 1,024 4-bit MDACs. The off-chip amplifiers for this preliminary system are provided by a pair of op amps run in the transimpedance mode, one each attached to the exciting and inhibiting lines. The outputs of the two amplifiers are then fed into a differential amplifier, which provides the output drive to other analog inputs.*

nal, starting at the upper left-hand corner, correspond to a match in all 16 bits. The variation in this maximum output is approximately five parts in 1,100 (decimal), or less than 0.5%. The boxed numbers correspond to a distance or mismatch of eight bits. Ideally, these values should all be the same; experimentally, they vary by less than 2%. When this experiment was originally performed, differences between nominally equivalent outputs showed much larger column-to-column variations than differences within a single column serviced by a single set of amplifiers. This effect was largely due to offset variations between amplifiers. The problem was

minimized by using low-offset amplifiers and by raising the input-voltage operating range to increase the signal-to-offset voltage ratio.

The Hamming net used single-bit weights. To demonstrate the 4-bit weight capability, a system was configured to act as a Gaussian classifier for a speech-recognition application. This network implements, in hardware, speech-recognition simulations that have been discussed in detail [3]. In this experiment an input vector that consisted of 22 continuous analog variables was fed through programmed weights (4-bit values stored in the MDACs) into each of seven summing nodes (the summing amplifiers). The design configuration is shown in Fig. 8. The inputs represent 11 cepstral components derived from two 10-m/s frames of speech samples. The speech samples were the seven numbers "one" to "seven." Reference 6 describes how the first stage, acting as a biased linear correlator, can effectively calculate the square of the Euclidean distance between the applied vector and each of the seven stored vectors.

The outputs from the first stage were each fed into one of seven nodes (differential amplifiers), which were configured to perform the winner-take-all function. All seven differential-amplifier outputs were fed to a single summing node, the output of which fed back to the differential amplifiers' inhibitory inputs. The negative excursions of the seven outputs were clamped nearly to ground by diodes. The clamping, along with the high gain used in the seven output nodes, caused the feedback voltage to track the largest output. Consequently, the other differential amplifiers' inputs were presented with an inhibiting voltage that clamped their output nodes to ground. The result was a circuit in which the output node with the greatest output (the winner) reproduced its positive input value and all other nodes produced very small voltages. To present a graphic demonstration of net operation, the clamp diodes on the winner-take-all, or max-picker, circuit were replaced with LEDs. The LEDs performed the clamping function, but also lit up on all losing nodes and remained unlit on the winning node.

The results of the experiment are shown in Table 2. Here 182 vectors were applied to the
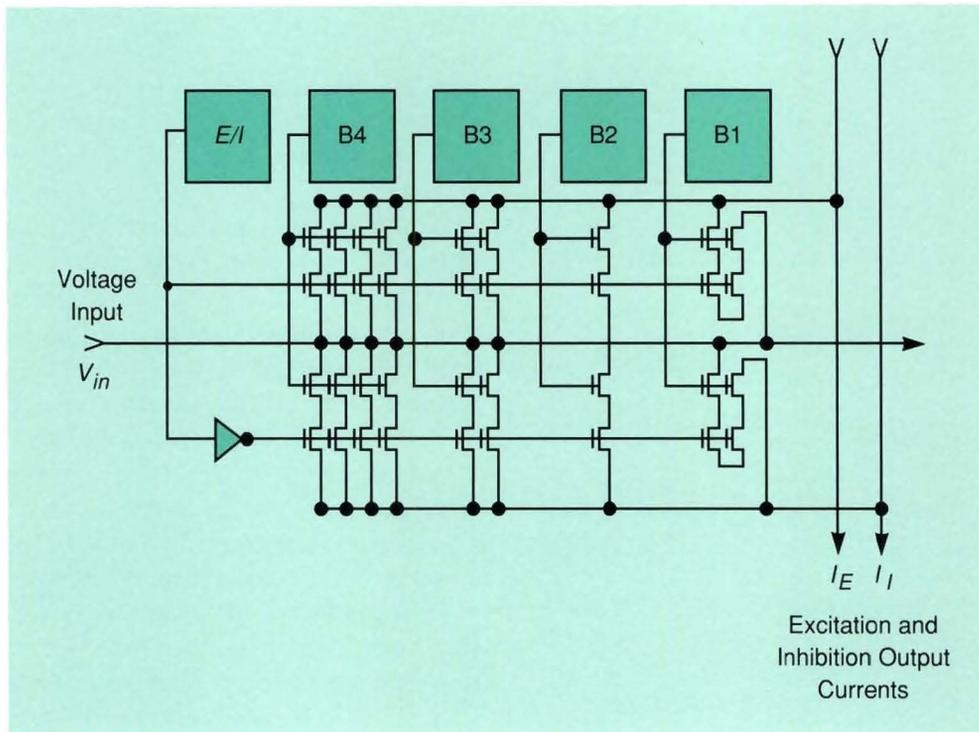
Fig. 5 — This MDAC circuit provides the weighted connections essential to neural-network operation. The weights are stored as digital values in bits 1 through 4; the E/I input is used to determine whether the output current is excitatory or inhibitory.

circuit and the outputs were compared with the results of a computer simulation that determined discrepancies between the hardware implementation and a mathematical simulation. Simulated and experimental winners are tabulated for each of the seven output nodes.

The numbers along the diagonals indicate the number of matches between the two categories. Off-diagonal components indicate misclassifications, of which there were seven out of the 182 trials. Five of these resulted from misidentifying type 1 vectors as belonging to type 5. This confusion is not too surprising, because vectors 1 and 5 are far closer than any other pair, although not so close that they should not have been distinguished by the circuit. This error is currently under investigation.

## Alternative-Circuit Realization

The MDAC circuit fabricated and used to implement the networks described above oper-
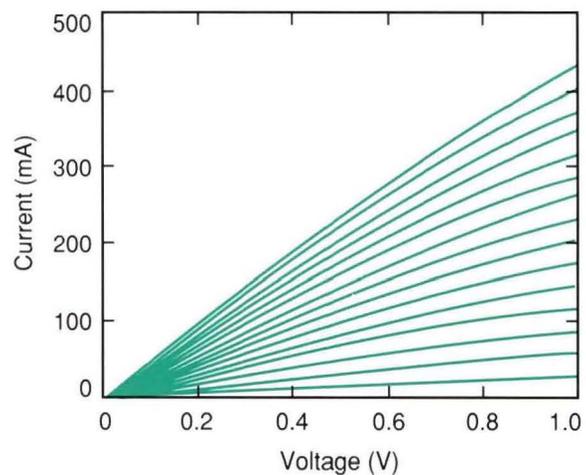


Fig. 6 — Output current of one of the MDACs used for weighted connections. The current steps along the vertical axis represent increments in the digital weight stored within the MDACs; the overall increase in output current exhibited along the horizontal axis is due to an increase in the voltage input to the MDAC. The current increments are uniform to within 5%.

INPUT VECTORS

Bias = +2

MATCH

1  1

+1
−1

(1010)  $1 \cdot 1 + 0 \cdot (-1) + 1 \cdot 1 + 0 \cdot (-1) + 2 = 4$

0  1

+1
−1

1  1

+1
+1

(1110)  $1 \cdot 1 + 1 \cdot 1 + 1 \cdot 1 + 0 \cdot (-1) + 1 = 4$

+1
−1

0  0

Bias = +1
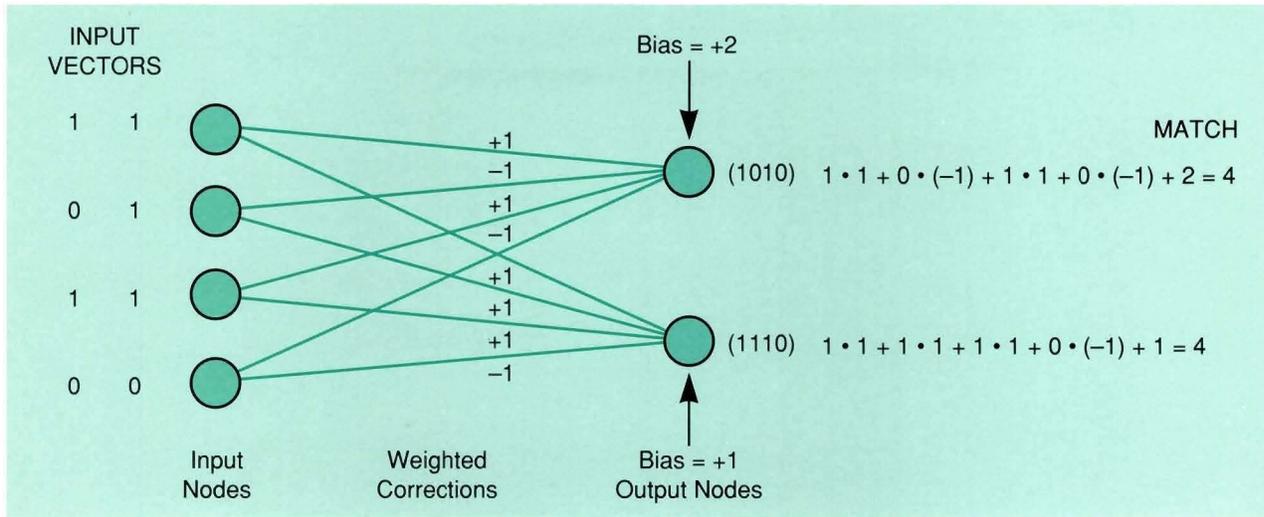
Input Nodes    Weighted Corrections    Output Nodes

Fig. 7 — This simple four-input two-output neural network illustrates the implementation of a neural-network model. The weights given as either +1 or −1 alongside the connections are effected by MDACs.

ates in the linear region of the transistor conductance curve for source-drain voltages up to 1 V. The percentage deviation from linearity varies directly with signal voltage in this range.

The MDAC's transistors can also be operated in saturation; then they perform the multiplication operation by driving the gate, rather than the drain. Using this saturation approach, we designed and fabricated a second version of the MDAC chip. In this chip, the subtraction of inhibiting current from exciting current uses current mirrors.

The MDAC shown at the top of Fig. 9 uses a Ones-complement representation to realize the weight strength and the connection polarity. Like the previous design, the storage bits B0 through B3 turn on transistors that correspond to the desired weight. For negative weights, B4 gates the lower row of transistors as well, forming the Ones-complement by exciting both inhibitory and excitatory lines. The analog input $V_{in}$ is applied to the second and third rows of transistors and generates, in each On transistor path, a current proportional to the output gen-

### Table 1. Results from the Hamming-Net Experiment

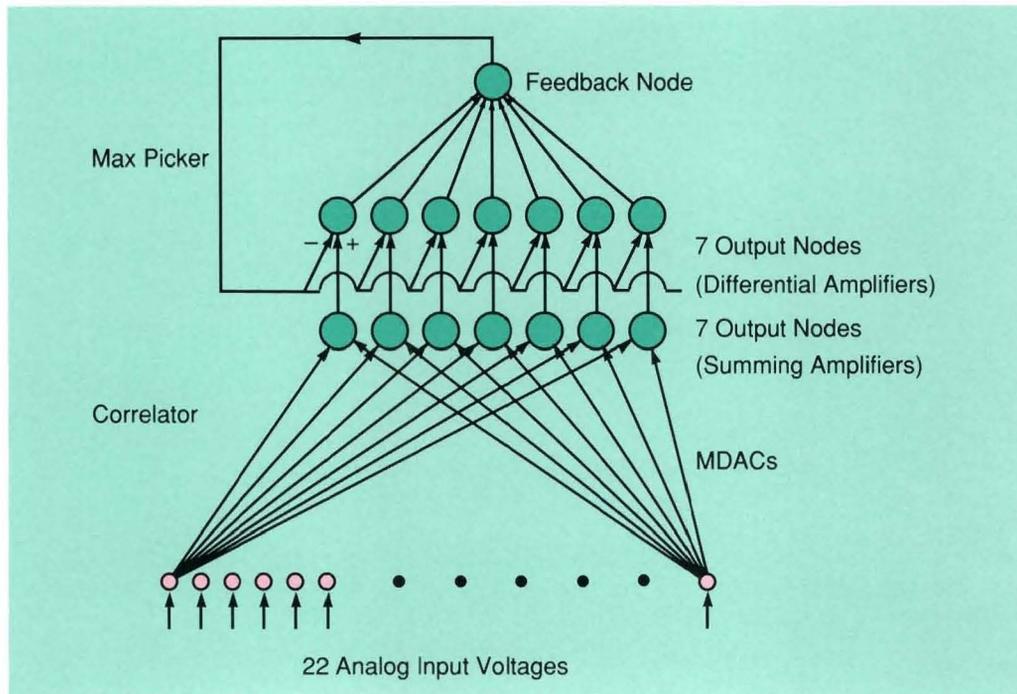| Applied Vector | Nearest Stored Vector | Outputs in Millivolts | | | | | |
|---|---|---|---|---|---|---|---|
| | | A | B | C | D | E | F |
| 1 | A | (1,130) | 695 | 556 | 556 | 775 | 300 |
| 2 | B | 710 | (1,125) | 554 | 700 | 636 | 559 |
| 3 | C | 556 | 560 | (1,125) | 555 | 637 | 560 |
| 4 | D | 556 | 695 | 558 | (1,125) | 493 | 845 |
| 5 | E | 780 | 631 | 629 | 485 | (1,130) | 353 |
| 6 | F | 303 | 556 | 557 | 840 | 358 | (1,130) |
| 7 | A | 1,065 | 628 | 480 | 627 | 705 | 352 |
| 8 | C | 428 | 556 | 705 | 557 | 640 | 560 |
| 9 | A&B | 705 | 695 | 557 | 415 | 493 | 300 |

Fig. 8 — This neural net consists of 154 MDACs connected to seven summing amplifiers. The outputs of the summing amplifiers are connected to seven differential amplifiers. The differential amplifiers and the feedback node comprise a max-picker circuit that finds the vector registering the strongest correlation with the input vector.

erated at another neural node.

One of the principal motivations for using current mirrors is to allow the direct connection of neuron outputs to inputs. Also, the high-impedance gate input is desirable for distributing signals in a wafer-scale device, because interconnection impedances are not negligible.

The cascode-mirror circuit shown at the bottom right of Fig. 9 subtracts the inhibition current from the excitation current. The current set up in the excitation branch is reflected in the other leg of the mirror. Any current in excess of the inhibition current is passed to ground through the lower transistor, which has a conductance that corresponds to the maximum MDAC weight. (The arrow indicates that the conductance of this transistor can be varied by programming a storage register, which controls a bank of variable-width transistors.) The gate of the lower transistor is maintained at the voltage necessary to cause the transistor to sink the required signal current, $I_E - I_I$, through the action of the differential amplifier whose output,

$V_{out}$, is the neural-node output voltage. If inhibition exceeds excitation, the diode connected to the inhibition leg prevents the internal node ($I_I$) from being pulled to ground.

The cascode-mirror arrangement was selected

**Table 2. Results of Gaussian-Classifier Experiment**

Winners by Network

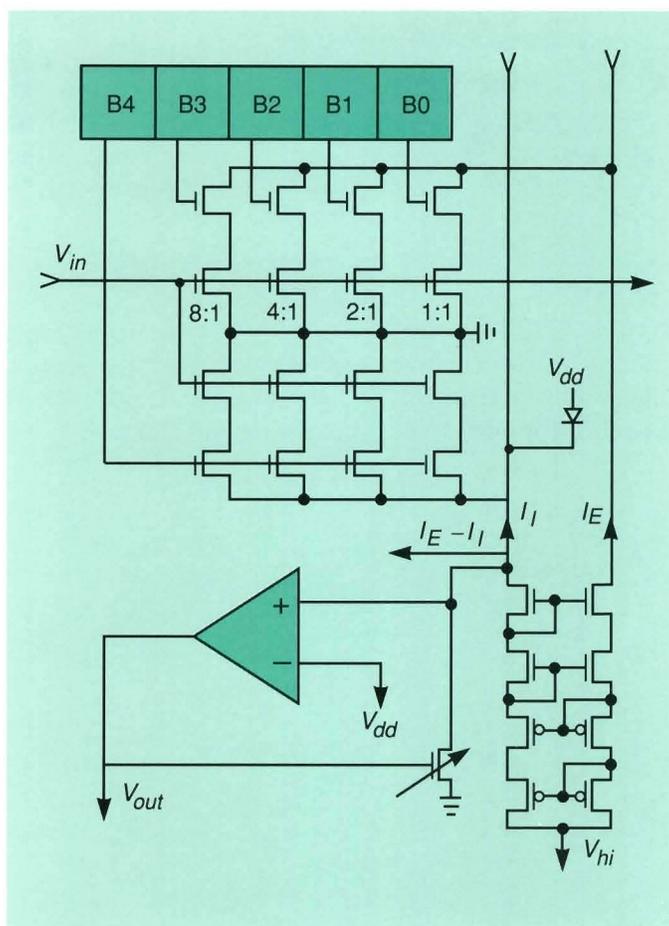| | | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|---|
| | 0 | 19 | | | | | | |
| | 1 | | 19 | | | | 5 | 1 |
| Winners by Simulation | 2 | | | 24 | | | | |
| | 3 | | | | 28 | | | |
| | 4 | 1 | | | | 30 | | |
| | 5 | | | | | | 29 | |
| | 6 | | | | | | | 26 |

*Fig. 9 — The current-mirror MDAC provides the high input imped-
ance required by a highly interconnected circuit.*

for its insensitivity to voltage variations at the
input and for its ability to hold all input ports
at the same potential. The equipotential input
ports guarantee that all the MDACs operate at
the same $V_{ds}$, which eliminates any effects due
to Early voltage. Because of the high imped-
ances in the circuit, however, the transistor
sizes necessary to handle any appreciable cur-
rents prohibit both a homogeneous distribution
of mirrors throughout the connection matrix
and the ability to accommodate the entire oper-
ating range of MDAC currents.

In this implementation, a single mirror at the
output of each neuron was used. The mirror was
designed to operate linearly over the range of
0 to 2 mA, which is nearly the output range of a
single MDAC. (The range of an MDAC is con-
strained by process geometry.) Operation of
the MDACs over a reduced range is accom-
plished by scaling the input/output mirrors
that set up the gate voltages distributed across
the chip. Through a 4-bit coefficient stored in
RAM, the conductances are variable with re-
spect to the full conductance of an MDAC in
ratios from 1:1 to 15:1.

## Conclusions

Results of experiments performed on small
networks built from integrated arrays of multi-
plying analog-to-digital converters controlled by
writable storage registers are encouraging.
These arrays implement the programmable
weights that are common to virtually all neuro-

morphic systems. The laser-linking technology and the design techniques that were used to build wafer-scale signal-processing systems are applicable to the design of very large monolithic neural networks. We estimate that a 4-in wafer, employing 2-$\mu$m design rules, will support a network consisting of about 10,000 summing nodes, with an average fan-in of 16. A network of this size will be able to implement hardware networks as large as the largest networks that have been simulated on general-purpose computers. Moreover, the wafer-scale circuit will improve computation rates by many orders of magnitude, and operate in microseconds, rather than seconds or minutes.

We have not yet addressed the development of on-wafer circuitry for adaptive weight modification or learning. This issue is now a major objective of our design efforts.

## Acknowledgments

# References

1. J.I. Raffel, A.H. Anderson, G.H. Chapman, K.H. Konkle, B. Mathur, and A.M. Soares, "A Wafer-Scale Digital Integrator Using Restructurable VLSI," *IEEE J. Solid State Circuits* **SC-20**, 399 (1985).

2. R.P. Lippmann, "An Introduction to Computing with Neural Nets," *IEEE ASSP Mag.* **4**, No. 2, 4 (1987).

3. R.P. Lippmann and B. Gold, "Neural-Net Classifiers Useful for Speech Recognition," *IEEE First Int. Conf. on Neural Networks, San Diego, Jun. 21–24, 1987*, p. 417; R.P. Lippmann, "Neural Network Classifiers for Speech Recognition," *Lincoln Laboratory Journal* **1**, 107 (1988).

4. S.L. Garverick and E.A. Pierce, "A Single Wafer 16-Point 16 MHZ FFT Processor," *IEEE Proc. of the 1983 Custom Integrated Circuits Conf., Rochester, NY, May 23–25, 1983*, p. 244.

5. J.R. Mann and F.M. Rhodes, "A Wafer Scale DTW Multiprocessor," *ICASSP 86 Proc., Tokyo, Apr. 7–11, 1986*, p. 1557.

6. D.E. Rumelhart and J.L. McClelland, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, 1986).

7. J. Alspector and R.B. Allen, "A Neuromatic VLSI Learning System," *Advanced Research in VLSI: Proc. of the 1987 Stanford Conf.* (MIT Press, Cambridge, MA, 1987), p. 313.

JACK I. RAFFEL is Leader of the Digital Integrated Circuits Group. He received an A.B. from Columbia College and a B.S. in electrical engineering from the Columbia School of Engineering. He received an M.S. degree in electrical engineering. Before joining Lincoln Laboratory, Jack was a research assistant at the Digital Computer Laboratory at MIT. His work has spanned the areas of ECL gate arrays, dielectric isolation, MNOS memory, high-speed analog/digital conversion, integrated-circuit CAD systems, and restructurable VLSI.

JAMES R. MANN is a staff member in the Digital Integrated Circuits Group. His research interests include computer architectures and digital and analog circuit design. Jim received B.S.E.E. and M.S.E.E. degrees from Michigan Technological University.

ROBERT BERGER is a staff member in the Digital Integrated Circuits Group. He received a B.E.E. from Rensselaer Polytechnic Institute, an M.S. in applied physics from Harvard University, and a Ph.D. in applied mathematics from Harvard. His current work focuses on wafer-scale integrated circuits. Bob has worked for Lincoln Laboratory for 28 years.

ANTONIO M. SOARES is a staff member in the Digital Integrated Circuits Group. He was born in Ponta Delgada, Azores, Portugal. Tony received a B.S. degree in electronics from Lowell Technological Institute. He joined Lincoln Laboratory in 1968 and now works on VLSI testing.

SHELDON L. GILBERT was a research assistant in the Digital Integrated Circuits Group, where he completed a master's thesis entitled, "Implementing Artificial Neural Networks in Integrated Circuitry: A Design Proposal for Back-Propagation." Prior to obtaining his master's degree in electrical engineering from MIT, he received a B.S. in electrical engineering from the University of Illinois. Sheldon is now employed by Spectrix (Evanston, IL), where he is developing data communication products.