Sigmoid neural transfer function realized by percolation

D. Pignon, P. J. M. Parmiter, J. K. Slack, M. A. Hands, and T. J. Hall

King's College, University of London, Strand, London WC2R 2LS, UK

M. van Daalen and J. Shawe-Taylor

Royal Holloway College, University of London, Egham, Surrey TW20 0EX, UK

Received September 12, 1995

An experiment using the phenomenon of percolation has been conducted to demonstrate the implementation of neural functionality (summing and sigmoid transfer). A simple analog approximation to digital percolation is implemented. The device consists of a piece of amorphous silicon with stochastic bit-stream optical inputs, in which a current percolating from one end to the other defines the neuron output, also in the form of a stochastic bit stream. Preliminary experimental results are presented. © 1996 Optical Society of America

Current interest in digital stochastic bit-stream representation of information for neural processing systems[1] opens up the possibility of using a simple analog percolation device to implement neural functionality. This is possible because of the emergent response of a finite percolation grid[2] to a matrix of stochastic inputs, which in turn renders such a device suitable for use in a free-space optoelectronic system.[3,4]

One can use a continuous stream of digital bits to represent an analog value by modulating the bits stochastically with a probability related to the value to be encoded. This results in a bit stream in which the proportion of 1's gives the probability p, where $p = x$ for an analog value x in the interval (0, 1) and $p = (x + 1)/2$ for x in the interval $(-1, 1)$. The advantage of such a representation is that simple digital hardware can be used, with the good error performance of digital systems combined with the functionality of analog processing.[1] For instance, multiplication of two bit streams—such as weight × input for a neuron—reduces to a simple bitwise XNOR [for the $(-1, 1)$ interval, or AND for the (0, 1) interval] of successive bit pairs from the two streams, producing a single bit stream representing the product of the two values.

In the context of an optoelectronic implementation[3,4] the high available bandwidth inherent in optical beamlet channels permits very high (gigahertz) bit-stream modulation rates.

The focus of the study presented here is to produce a simple device to perform the summing and thresholding functions of a neuron.[5] We are working on a three-dimensional optoelectronic implementation of a neural network[6] that uses planes of simple electronic devices for processing, densely interconnected by optical beamlet channels. Bearing in mind such a three-dimensional optoelectronic implementation, what is required is a planar device to receive a two-dimensional grid of optical bit-stream inputs (already multipled by their corresponding weights by an array of optoelectonic XNOR gates in the preceding plane), to sum them, and then to apply a sigmoid transfer function to the sum, thus providing the single analog output from the neuron in stochastic bit-stream form.

This implementation may be simply realized by a two-dimensional network of percolating cells (Fig. 1).[2] Each input is incident upon a single cell in the grid (of $N \times N$ cells) that then either conducts or not, depending on whether the incident optical bit is high or low. Each cell is connected to its nearest neighbors (four connected) or further [eight connected (see Fig. 2) etc.], and a voltage is applied across the top and bottom rows of cells. At any one instant the collective bits from the incident streams may or may not cause a conducting path between the contacts. If the current flowing between the contacts is interpreted digitally, i.e., no current = 0 and some current = 1, another stochastic bit stream results that represents the sum of the inputs with a sigmoidlike transfer function applied, as desired. The sigmoidlike shape of the transfer function appears as an emergent property of percolation through a finite grid.[7] The smaller the grid, the shallower the curve, and in the limit of an infinite grid the curve becomes a step. Also, a higher connectivity makes the curve steeper. Figure 3 shows a computer simulation of the bit-stream digital percolation grid.

We report on an experiment to assess the feasibility of further simplification of the device by the use of an analog approximation. This consists of a two-dimensional piece of amorphous silicon with the optical bit-stream inputs incident upon the face and contacts appended on two sides as described above

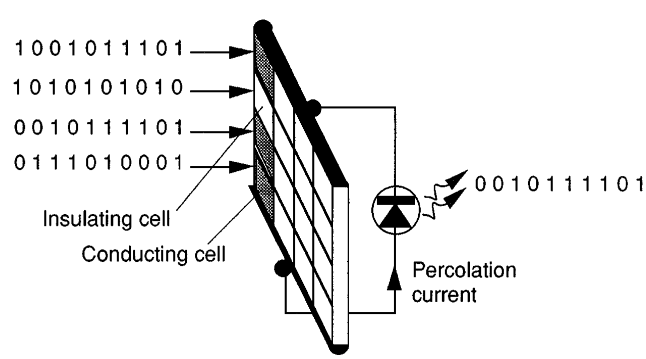

Fig. 1. Digital percolation grid with bit-stream inputs and outputs.

0146-9592/96/030222-03$6.00/0

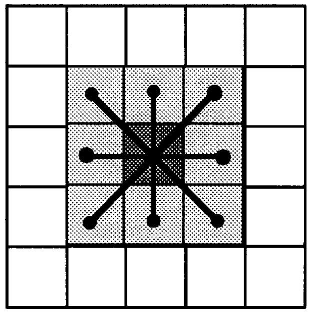

Fig. 2. Example of an eight-connected cell.

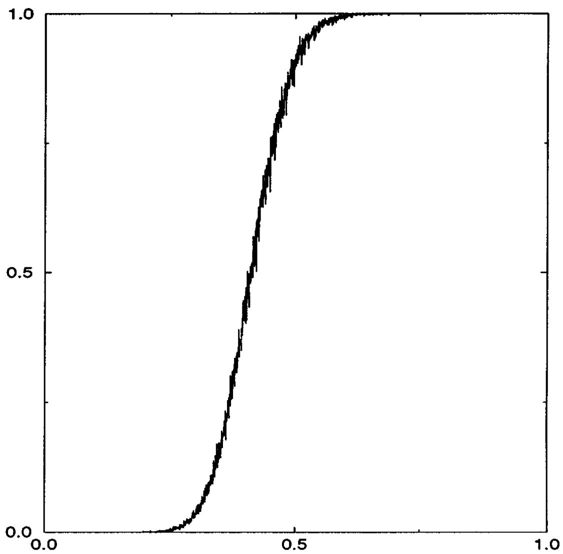

Fig. 3. Computer simulation of a digital percolation grid. 16 × 16, 8 connected, 100 kbits per input.

(Fig. 4). Each cell is now defined by the input beamlets themselves, where a high bit (high intensity) produces photogenerated charge carriers in its cell (of resistivity R_{on}) and a low bit (low intensity) leaves a cell of high resistivity, R_{off}. A rough calculation suggests that for this device to be equivalent to the digital one described above the necessary optical contrast ratio between high and low states is given by

$$R_{\mathrm{off}}/R_{\mathrm{on}} \geq N(N-2)+1\,,$$

with the absolute optical power in each channel scaling in a similar manner because of the need to overcome background conductivity of the silicon.

In the experiment we produced the bit streams by modulating an expanded laser beam, using a ferroelectric liquid-crystal spatial light modulator (SLM). The input beam was chopped, and a lock-in amplifier was used to measure the current across the device. A computer was used to produce the bit streams on the SLM and to gather the data. After a threshold current was determined by means of two test patterns, a current above this threshold produced a 1 output bit, and a current below produced a 0. We attached probe wires to the amorphous silicon wafer, using silver DAG paint, creating ohmic, rather than Shottky as would be expected from a clean metal–semiconductor junction, contacts, as required.

Initially we used a 5 × 5 grid to obtain a shallow curve and also because, as mentioned above, this would require a contrast ratio of only 16:1 between the high and the low signals to be equivalent to the digital percolation as simulated on the computer. However, with the particular setup used, because of leakage between SLM pixels and reflections, the contrast ratio obtained was 14:1.

Nevertheless, preliminary results are highly encouraging, as shown by Fig. 5. The curve was found to be almost exactly described by the sigmoid form

$$y = 1/[1+\exp(ax+b)]\,,$$

with coefficients a and b equal to -20.0 and 8.9, respectively. One can translate the curve along the x axis (b modified) at will by varying the current threshold.

At a contrast ratio of 16:1 or greater the device operates as a digital percolation matrix with a connectivity that depends on the contrast ratio; the lower the ratio, the more effect a current through a distant cell will have on the cell in question, increasing the effective connectivity. At a ratio below this value (as indeed was the case with the preliminary runs) a second, analog, effect begins to affect the result. This is the position-independent sum of cell resistivities. In the experiment the summing effect alone would

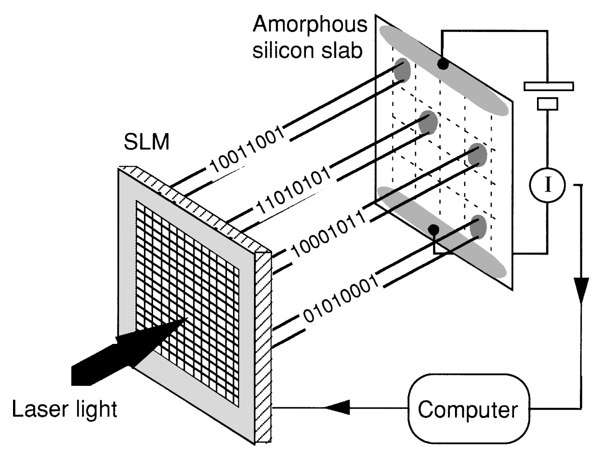

Fig. 4. Schematic of the analog percolation experiment. For clarity, the imaging optics are not shown.

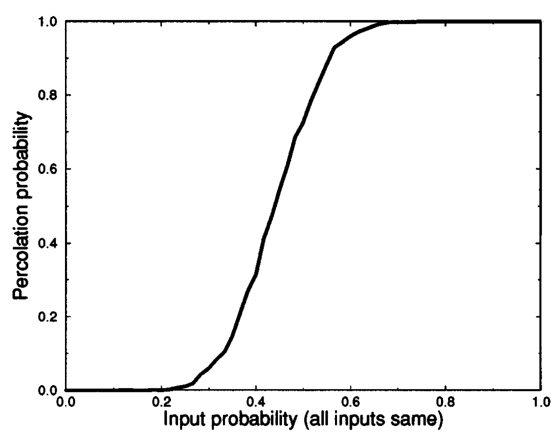

Fig. 5. Experimental result. 5 × 5 grid, 1500 bits, 60 steps.

produce a step. In reality there is a combination of summing and percolation, the relative importance of each being determined by contrast ratio. The effect of the summing on the percolation graph is to make it steeper.

In summary, we have shown that an analo opto-electronic percolation grid can be used to perform the sigmoid transfer function required by a neuron. Even from a preliminary nonoptimized experiment the result obtained was highly encouraging. Further quick improvements under way, such as the introduction of a pixelating computer-generated hologram, should result in an improvement of contrast ratio to 100:1, permitting a percolation grid of 11×11 or even more cells. Methods for determining the rate of change of the output with respect to each input already exist for the perfect digital case, so these will also be investigated. Gradient information is helpful to neural network training algorithms.

The advantage of such a device is its great ease of fabrication and low cost, because its natural emergent properties are exploited without the need for fabrication of complex devices.

The research of D. Pignon is supported by an Engineering and Physical Sciences Research Council Ph.D. Cooperative Award in Science and Engineering cosponsored by GEC-Marconi. The authors thank Bill Crossland and Steven Latham, then at Cambridge University, for supplying the amorphous silicon wafer for use in this experiment, and Pilkingtons, British Aerospace, and GEC for providing other equipment.

References

1. J. Shawe-Taylor, P. Jeavons, and M. van Daalen, Connect. Sci. **3,** 317 (1991).
2. T. J. Hall and M. van Daalen, UK patent application 9420449.2 (October 11, 1994).
3. D. Pignon and T. J. Hall, “A quantised optical neural network with on-line learning,” submitted to Int. J. Optoelectron.
4. A. G. Kirk, S. Jamieson, H. Imam, and T. J. Hall, Opt. Comput. Process. **2,** 293 (1992).
5. P. K. Simpson, *Artificial Neural Systems* (Pergamon, Oxford, 1990).
6. T. J. Hall, W. Peiffer, M. Hands, H. Thienpont, W. A. Crossland, J. S. Shawe-Taylor, and M. van Daalen, in *Optical Computing,* Vol. 10 of 1995 OSA Technical Digest Series (Optical Society of America, Washington D.C., 1995), p. 210.
7. D. Stauffer and A. Aharony, *Introduction to Percolation Theory* (Taylor & Francis, London, 1992).