FLOATING-GATE DEVICES: THEY ARE NOT JUST FOR DIGITAL MEMORIES ANYMORE

Paul Hasler[1], Bradley A. Minch[2], and Chris Diorio[3]

ABSTRACT

Since the first reported floating-gate structure in 1967, floating-gate transistors have been used widely to store digital information for long periods in structures such as EPROMs and EEPROMs. Recently, floating-gate devices have found applications as analog memories, analog and digital circuit elements, and adaptive processing elements. Floating-gate devices have found commerical applications, e.g. ISD, for long-term non-volatile information storage devices for analog applications. The focus of floating-gate devices has been towards fabrication in standard CMOS processes, as opposed to the specialized processes for fabricating digital nonvolatile memories. Floating-gate circuits can be designed at any or all of three levels: analog memory elements, capacitive-based circuit elements, and adaptive circuit elements.

In 1967, Kahng and Sze reported the first floating-gate structure as a mechanism for nonvolatile information storage [1]. Since then, floating-gate transistors have been used widely to store digital information for long periods in structures such as EPROMs, EEPROMs, and Flash memories [2, 3]. Figure 1 sketches the growth of non-volatile digital technology from the inception of the floating-gate device in 1967. These digital nonvolatile memory technologies have only been fabricated on specialized IC processes. Even though digital memories have been the primary application for floating-gate devices, recently floating-gate devices have been used as circuit elements. Floating-gate devices are not inherently digital memories, it only depends upon the way the circuit designer uses the available technology.

Figure 1 also sketches the progress of floating-gate devices and circuits other than EEPROM devices. The research in floating-gate circuits has been exponentially increasing over the last several years; the goal is to further develop this technology to give researchers a wider set of solutions in solving integrated circuit problems. The focus of this paper, which introduces the special session on floating-gate devices and circuits, is to show that floating-gate devices are not just for memories anymore, but are circuit elements with analog memory and important time-domain dynamics.

[1] Georgia Institute of Technology, Atlanta, GA 30332-0250, e-mail: phasler@ee.gatech.edu

[2] Cornell University, Ithaca, NY 14853-5401, e-mail: minch@anise.ee.cornell.edu

[3] University of Washington, Seattle, WA 98195-2350, e-mail: diorio@cs.uwashington.edu

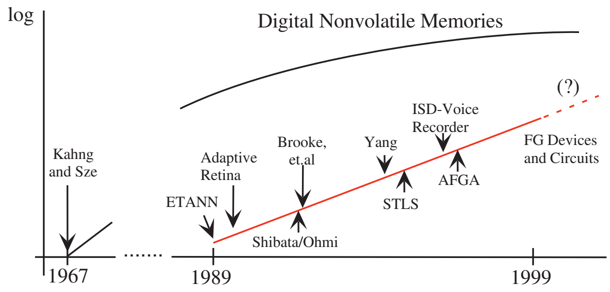

Figure 1: Circuit diagram of progress in floating-gate devices. In the overall picture, digital memories (EPROMs, EEPROMs and Flash memories) have been the primary application for floating-gate elements, and have been a very important part of today's market. Floating-gate circuits, although currently a small part of the industry, have been rapidly increasing over the last 10 years.

1. INITIAL WORK ON FLOATING-GATE CIRCUITS

The current interest in floating-gate circuits started from developing large-scale computations in neuromorphic systems; several examples of neuromorphic systems are described elsewhere [4]. The introduction of the ETANN chip in 1989 showed the potential of using floating-gate devices for applications other than digital memory elements [5]. The core of the chip's function was an analog vector-matrix calculation employing a floating-gate array of 10240 analog matrix elements. The floating-gate devices, developed in a specialized EEPROM fabrication process, were used as an analog floating-gate memory; the output of these cells were currents that were directly used by multiplier circuits. Furthermore, this Intel product was considered for various applications, and the resulting insights would be crucial in building future floating-gate chips, insights that would not have been demonstrated in any other way. One issue was that the interfacing, packaging, and cost of the ETANN chip often offset the impressive computational ability of the chip.

Three research accomplishments laid the groundwork for much of the current floating-gate circuit development. First, Thomsen and Brooke's demonstration and use of electron tunneling in a standard CMOS double-poly process allowed many researchers to investigate floating-gate circuit concepts without requiring access to specialized fabrication processes [6]. Second, the νMOS, or neuron-MOS, circuit approach by Shibata and Ohmi provided the initial inspiration and framework to use capacitors for linear

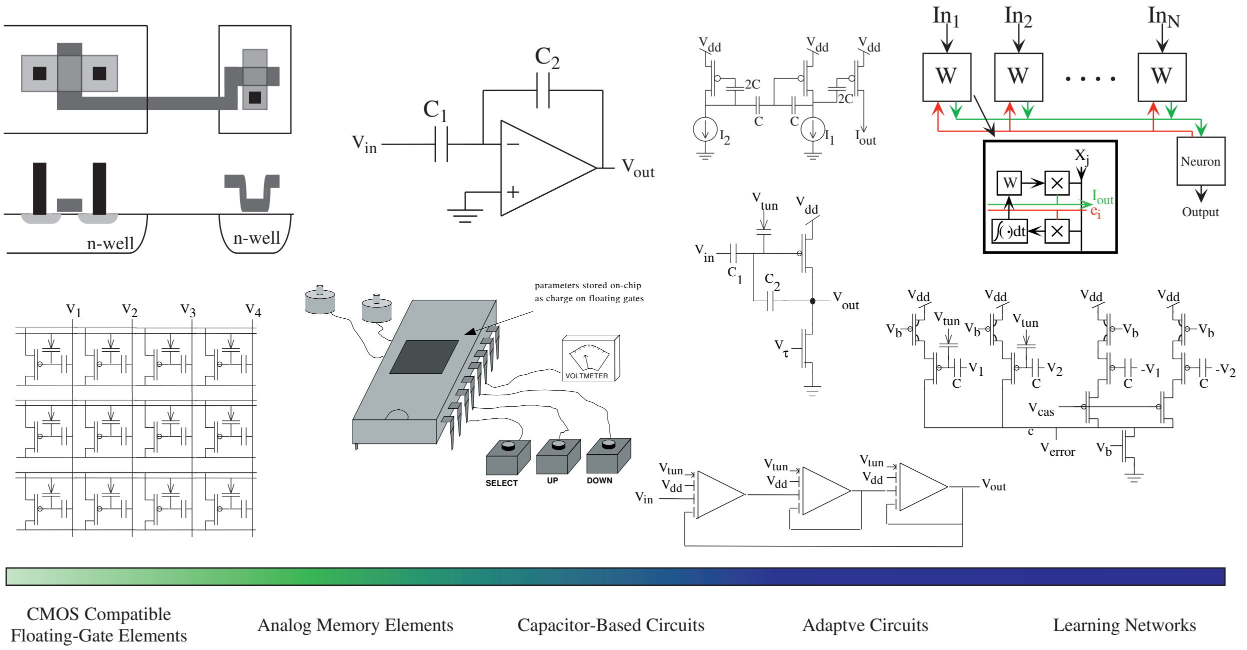

Figure 2: Spectrum of floating-gate device and circuit research, with representatives of floating-gate circuits and applications. Several more quality examples can be found in the papers given at this session. [9]

computations [7]. These researchers concentrated on the floating-gate circuit properties instead of the device properties, and used either ultra-violet (UV) light to equalize charge, or simulated floating-gate elements by opening and closing MOSFET switches. Third, Carver Mead's adaptive retina gave the first example of using contiuously-operating floating-gate programming/erasing techniques, in this case UV light, as the backbone of an adaptive circuit technology [8].

2. CURRENT STATE OF FLOATING-GATE DEVICE AND CIRCUITS

The last six years have seen an explosion of floating-gate technology applied not only to non-volatile memories, but also towards floating-gate circuits. Although neuromorphic applications have been a prominent use of this technology, the focus has shifted towards utilizing floating-gate circuits to solve practical analog and digital circuit problems. This explosion has resulted in a wide range of devices and circuits that can be fabricated in either a single-poly or a double-poly process.

Figure 2 shows the spectrum of floating-gate device and circuit research, with a representative selection of circuit illustrations of particular floating-gate circuits and applications. Several more quality examples can be found in the papers given at this session [9]. In the following three subsections, we will briefly overview the three major thrusts in floating-gate device and circuits. First, floating-gate devices are used as analog memory elements. Second, floating-gate devices are used as part of capacitive-based circuits. Third, floating-gate devices are used as adaptive circuit elements. We have written detailed overviews of these topics elsewhere [10, 11, 12, 13, 14].

2.1. CMOS Floating-Gate Memory Elements

Floating-gate devices have been used for long-term non-volatile information storage devices for analog applications. Since digital storage is the primary application of floating-gate devices, the next extension was to develop analog floating-gate memories. There is nothing inherently digital in the use of EEPROMs except in the implementation of the reading and writing circuitry. The development of programmable analog floating-gate memory arrays [15] eventually resulted in ISD's analog EEPROM chips that have found applications in audio recording applications [16].

A floating gate is a polysilicon gate surrounded by SiO_2. Charge on the floating gate is stored permanently, providing a long-term memory, because it is completely surrounded by a high-quality insulator. Furthermore, the charge on this floating-gate can be modified by projecting UV light on the chip, by applying large voltages across a silicon-oxide capacitor to tunnel electrons though the oxide, or by adding electrons using hot-electron injection. The physical effects of hot-electron injection and electron tunneling become more pronounced as the line-widths of existing processes are further scaled down [17], improving our floating-gate circuits.

Harrison and colleagues used floating-gate technology to eliminate off-chip-biasing voltages in the existing system by providing these voltages on-chip with arrays of programmable floating-gate voltages [18]. Modern analog, neuromorphic, or mixed-mode VLSI chips typically have large numbers of inputs and analog parameters, and the number of available pins is often a limiting factor in these systems. By moving the analog parameters and circuit biases onto the chip, we eliminate many pins that would normally connect to external potentiometers. The array of floating-gate memory elements can be individually programmed either up or down by straightforward digital controls. This chip uses an active capacitor compensation scheme to null out the capacitive effects resulting from changing the tunneling or drain voltage; practical system design often must take these effects into consideration.

2.2. Computation with floating-gate devices

Shibata and Ohmi called their floating-gate devices neuron-MOS transistors [7], based on the loose analogy between the function performed by these devices and by cells in the nervous system. Yang and Andreou refer to such devices as multiple-input floating-gate MOS (FGMOS) transistors [19]. Ramirez-Angulo calls them multiple-input floating-gate (MIFG) transistors [20]. The way of thinking about FGMOS transistors introduced by Shibata and Ohmi has resulted in a number of interesting analog and and digital information-pocessing circuits, including a multiple-input floating-gate differential amplifier [19], a four-quadrant floating-gate multiplier [20], simple D/A converters [21], and translinear circuits [22].

The floating-gate voltage, determined by the charge stored on the floating-gate, can modulate a channel between a source and drain, and therefore can be used in computation. Floating-gate circuits provide IC designers with a practical a capacitor-based technology; capacitors, rather than resistors, are a natural result of a MOS process. Floating-gate devices can compute a wide range of translinear functions by a particular choice of capacitive couplings into floating-gate devices [22]. Minch generalized the concept of translinear circuits using the inspiration resulting from subthreshold MOS transistors and capacitive voltage dividers [13]. Later in this session, Minch will present floating-gate log-domain circuits [9], Ramirez-Angulo will present his floating-gate circuit work, and Harrison et. al. will present floating-gate circuits used in the biological modeling of fly vision.

A second application of this technology is to build circuits with a programmable threshold voltage; therefore the threshold voltage can be set to a convienant location for a given circuit. Berg and Lande have developed a generalized floating-gate technique for tuning MOS-transistor circuits for low voltage (less than 1V power supply) operation [9, 23]. These researchers program the threshold voltages of both pFETs and nFETs, fabricated in a basic double-poly CMOS processes, using UV-light and reverse-biased power rails.

2.3. Adaptive floating-gate devices

The most novel and exciting development involving floating-gate devices are circuits have the inherent ability to adapt to the incoming and outgoing signals. Once the adaptation is finished, the resulting network state is preserved nearly indefinitely, due to the nonvolatile property of floating-gate devices. This property was the fundamental motivation behind the development of single-transistor synapses, single floating-gate FETs that emulate some of the computational and adaptive properties of biological synaptic elements [24].

The single-transistor synapse resulted from two circuit innovations [10, 12, 24, 25]. First, interpoly electron tunneling, used in early floating-gate applications [6], is unacceptable for continuously adapting applications due to the rapid oxide degradation. Electron tunneling through a MOS capacitor and pFET hot-electron injection are significantly more reliable schemes; the Etox EEPROM cell, developed independently [3], also uses electron tunneling though gate oxide for increased reliability. Second, the floating-gate currents must be modulated in a continuous fashion without significantly altering the circuit's behavior; therefore, slow updates can occur simultaneously with circuit computations. This thinking was the primary reason for choosing a complementary form of electron tunneling and hot-electron injection to modify the floating-gate charge.

Of the several examples of continuously-adapting floating-gate circuits [10], one of the most elegant examples of an adaptive floating-gate pFET circuit is the autozeroing floating-gate amplifier (AFGA) [10, 26]. The AFGA uses tunneling and pFET hot-electron injection to adaptively set its DC operating point. The modulation of the pFET hot-electron injection by the output voltage provides the correct feedback to return the output voltage to the proper operating regime. Because of feedback applied to the floating gate, this adaptation is an inherent part of the circuit's operation—no additional control circuitry is required. The AFGA demonstrates how to use continuous-time, floating-gate adaptation in amplifier design, and is an example of how one of many classical engineering problems is solvable using floating-gate techniques. AFGAs have been the basis of many capacitor-based continuous-time filters and sensor interface circuits [10]. This continuous adaptation shows a method to take advantage of very small devices that, under normal operation, have non-negligible hot-electron and tunneling currents [27].

Recently, building networks of continuously adapting floating-gate synapses has received considerable interest. Current results clearly show that floating-gate synapses, and the resulting circuits, adapt to encode the statistics of incoming input signals [9, 28]. Hafliger and Rasche will present an analog floating-gate model of biologically realistic 'silicon neurons' using floating-gate learning synapses, that perform on-chip learning by emulating a form of long-term potentiation (LTP) and long-term depression (LTD) as observed in biological neurons [9].

3. FUTURE OF FLOATING-GATE CIRCUITS

The future looks promising for floating-gate circuits. For floating-gate circuits to continue on its ever-expanding trend, a few achievable hurdles must be overcome. First, the floating-gate charge must be quickly modified during initial test proceedures to move all the floating-gate circuits to a useful operating range. Most floating-gate circuits exhibit very robust behaviors once they are in the neighborhod of their proper operating range. Typical constraints for industrial IC testing require characterization times in the millisecond range, a requirement not satisfied by most current floating-gate circuits. These constraints will be more significant for chips utilizing hundreds or thousands of floating-gate devices, as is expected with the adaptive floating-gate applications. A second issue is to increase the industrial applications of floating-gate devices, and to look for applications to imbed these circuits into various system applications. A final issue is to strengthen the community of researchers working in this field, to encourage new researchers in this field, and to build a common language to facilitate collaborations in this area.

4. REFERENCES

[1] D. Kahng and S.M. Sze, "A floating-gate and its application to memory devices," The Bell System Technical Journal, vol. 46, no. 4, 1967, pp. 1288-1295.

[2] F. Masuoka, R. Shirota, and K. Sakui, "Reviews and prospects of non-volatile semiconductor memories," IEICE transactions, vol. E 74, no. 4, 1991, pp. 868-874.

[3] S. Lai, "Flash memories: where we were and where we are going," IEEE International Electron Devices Meeting, San Francisco, 1998, pp. 971-974.

[4] C. Mead, Analog VLSI and Neural Systems, Addison-Wesley, Reading, MA, 1989.

[5] M. Holler, S. Tam, H. Castro, and R. Benson, "An electrically trainable artificial neural network with 10240 'floating gate' synapses," Proceedings of the International Joint Conference on Neural Networks, Washington, D.C., vol. II, 1989, pp. 191-196.

[6] A. Thomsen and M.A. Brooke, "A floating gate MOSFET with tunneling injector fabricated using a standard double-polysilicon CMOS process," IEEE Electron Device Letters, vol. 12, 1991, pp. 111-113.

[7] T. Shibata and T. Ohmi, "A functional MOS transistor featuring gate-level weighted sum and threshold operations," IEEE Transactions on Electron Devices, vol. 39, no. 6, 1992, pp. 1444-1455.

[8] C.A. Mead and M. Ismail, editors, Analog VLSI Implementation of Neural Systems, Kluwer Academic Publishers, Norwell, MA, 1989.

[9] Special Session on Floating-Gate Devices and Circuits, International Symposium on Circuits and Systems, Orlando, 1999.

[10] P. Hasler, B. A. Minch, and C. Diorio, "Floating-Gate Devices and Circuits," in Advanced Research in VLSI, 1999. Also at http://users.ece.gatech.edu/phasler.

[11] P. Hasler, Foundations of Learning in Analog VLSI, Ph.D. Thesis, California Institute of Technology, February 1997. Also at http://users.ece.gatech.edu/phasler.

[12] C. Diorio, P. Hasler, B.A. Minch, and Carver A. Mead, "Floating-gate MOS synapse transistors," in Tor Sverre Lande, editor, Neurmorphic Systems Engineering: Neural Networks in Silicon, Kluwer Academic Publishers, Norwell, MA, 1998, pp. 315-338.

[13] B. A. Minch, Analysis, Synthesis, and Implementation of Networks of Multiple-Input Translinear Elements, Ph.D. Thesis, California Institute of Technology, May 1997.

[14] C. Diorio, Neurally Inspired Silicon Learning: From Synapse Transistors to Learning Arrays, Ph.D. Thesis, California Institute of Technology, May 1997.

[15] H.V. Tran, T. Blyth, D. Sowards, L. Engh, B.S. Nataraj, T. Dunne, H. Wong, V. Serin, T. Lam, H. Hazarian, and G. Hu, "A 2.5V 256-level non-volatile analog storage device using EEPROM technology," Proceedings of IEEE International Solid-State Circuits Conference, 1996, pp. 270-271.

[16] ISD Data Book: Voice Recording and Playback ICs, ISD, San Jose, 1996, 2nd edition.

[17] C. Mead, "Scaling of MOS technology to submicrometer feature sizes," Journal of VLSI Signal Processing, vol. 8, 1994, pp. 9-25.

[18] R.R. Harison, P. Hasler, and B.A. Minch, "Floating-gate CMOS analog memory cell array," Proceedings of IEEE International Symposium on Circuits and Systems, Monterey, 1998. Also at http://users.ece.gatech.edu/phasler.

[19] K. Yang and A.G. Andreou, "A Multiple Input Differential Amplifier Based on Charge Sharing on a Floating-Gate MOSFET," Journal of Analog Integrated Circuits and Signal Processing, vol. 6, no. 3, 1994, pp 197-208.

[20] J. Ramirez-Angulo, "+-0.75V BiCMOS Four Quadrant Analog Multiplier with Rail-Rail Input Signal-Swing," Proceedings of the 1996 IEEE International Symposium on Circuits and Systems,Atlanta, GA, vol. 1, 1996, pp. 242-245.

[21] T. Shibata and T. Ohmi, "Neuron MOS Binary-Logic Integrated Circuits—Part I:Design Fundamentals and Soft-Hardware-Logic Circuit Implementation," IEEE Transactions on Electron Devices, vol. 40, no. 3, pp. 570-576, 1993.

[22] B. A. Minch, C. Diorio, P. Hasler, and C.A. Mead, "Translinear circuits using subthreshold floating-gate MOS transistors," Analog Integrated Circuits and Signal Processing, vol. 9, no. 2, 1996, pp. 167-179.

[23] Y. Berger, D.T. Wisland, T.S. Lande, and S. Mikkelsen, "Ultra low-voltage digital floating-gate UVMOS Circuits" Proceedings of IEEE International Symposium on Circuits and Systems, Monterey, 1998.

[24] P. Hasler, C. Diorio, B.A. Minch, and C. Mead, "Single transistor learning synapses," in Advances in Neural Information Processing Systems 7, MIT Press, Cambridge, MA, 1995, pp. 817-824. Also at http://users.ece.gatech.edu/phasler.

[25] C. Diorio, P. Hasler, B.A. Minch, and C. Mead, "A single-transistor silicon synapse," IEEE Transactions on Electron Devices, vol. 43, no. 11, 1996, pp. 1972-1980.

[26] P. Hasler, B.A. Minch, C. Diorio, and C. Mead, "An autozeroing amplifier using pFET hot-electron injection," Proceedings of the International Symposium on Circuits and Systems, Atlanta, vol. 3, 1996, pp. 325-328.

[27] G. Timp, K.K. Bourdelle, J.F. Bower, F.H. Baumann, T. Boone, R. Cirelli, K. Evans-Lutterodt, J. Garno, A. Ghetti, H. Grossmann, M. Green, D. Jacobson, Y. Kim, R. Kliman, F. Klemens, A. Kornblit, C Lochstampfor, W. Mansfield, S. Moccio, D.A. Muller, L.E. Ocola, M.L. O'Malley, J.Rosamilia, J. Sapjeta, P. Silverman, T. Sorsch, D.M. Tennant, W. Timp, B.F. Weir, "Progress toward 10nm CMOS devices," IEEE International Electron Devices Meeting, San Francisco, 1998, pp. 615-618.

[28] P. Hasler and J. Dugger, "Correlation Learning Rule in Floating-Gate pFET Synapses," International Symposium on Circuits and Systems, Orlando, 1999.