FLOATING-GATE CMOS ANALOG MEMORY CELL ARRAY

Reid R. Harrison[1] *Paul Hasler*[2] *Bradley A. Minch*[3]

[1]Computation and Neural Systems Program 139-74, California Institute of Technology, Pasadena, CA 91125
harrison@klab.caltech.edu
[2]Department of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250
[3]School of Electrical Engineering, Cornell University, Ithaca, NY 14853

ABSTRACT

The complexity of analog VLSI systems is often limited by the number of pins on a chip rather than by the die area. Currently, many analog parameters and biases are stored off chip. Moving parameter storage on chip could save pins and allow us to create complex programmable analog systems. In this paper, we present a design for an on-chip non-volatile analog memory cell that can be configured in addressable arrays and programmed easily. We use floating-gate MOS transistors to store charge, and we use the processes of tunneling and hot-electron injection to program values. We achieve greater than 13-bit precision with no crosstalk between memory cells.

1. INTRODUCTION

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. Often, these parameters and other biases are stored off chip as voltages programmed by potentiometers or other sources, one pin per variable (see Figure 1a). Pins are also needed for inputs and outputs. The number of pins available on a chip is limited by the perimeter of the chip, which grows as the square root of die area. If we could move the analog parameters and circuit biases onto the chip, we would and save more pins for input, output, and diagnositics (see Figure 1b).

If we think of potentiometers as easily-modifiable voltage sources, then we would like to build an electronic version of the pot, an **e-pot**, that can be placed on the chip itself. If large numbers of these cells can be placed on one chip, then they will supply more biases than could the few pins required to program them. Also, storing biases on chip could significantly reduce circuit board complexity. We have the following requirements for our e-pot:

1. Non-volatile: Each e-pot must maintain its state indefinitely in the absence of power.

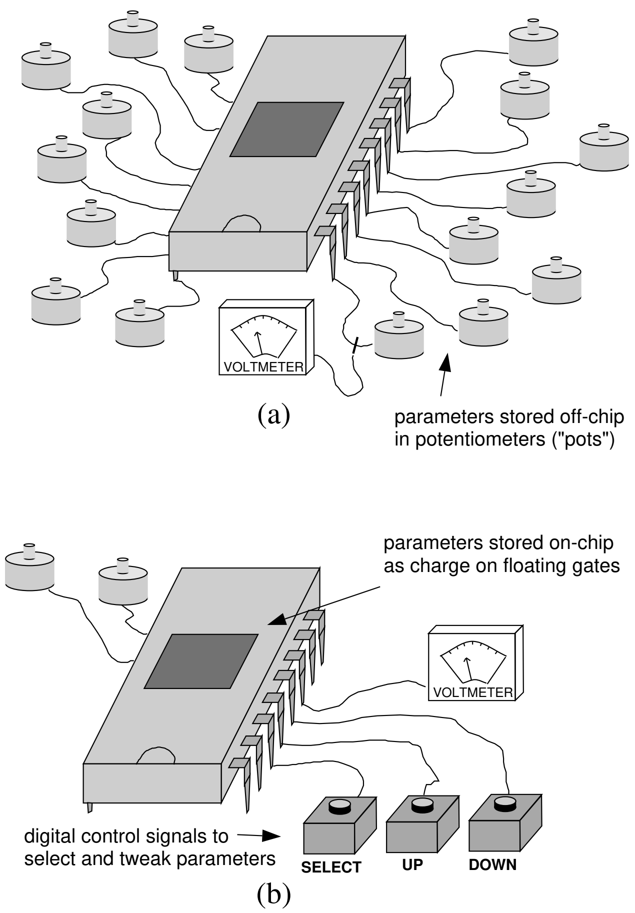

Figure 1. Using electronic potentiometers (e-pots). (a) On typical analog VLSI chips, many of the pins are consumed by bias voltages set with off-chip potentiometers. (b) By storing these voltages on chip, many pins are freed for I/O.

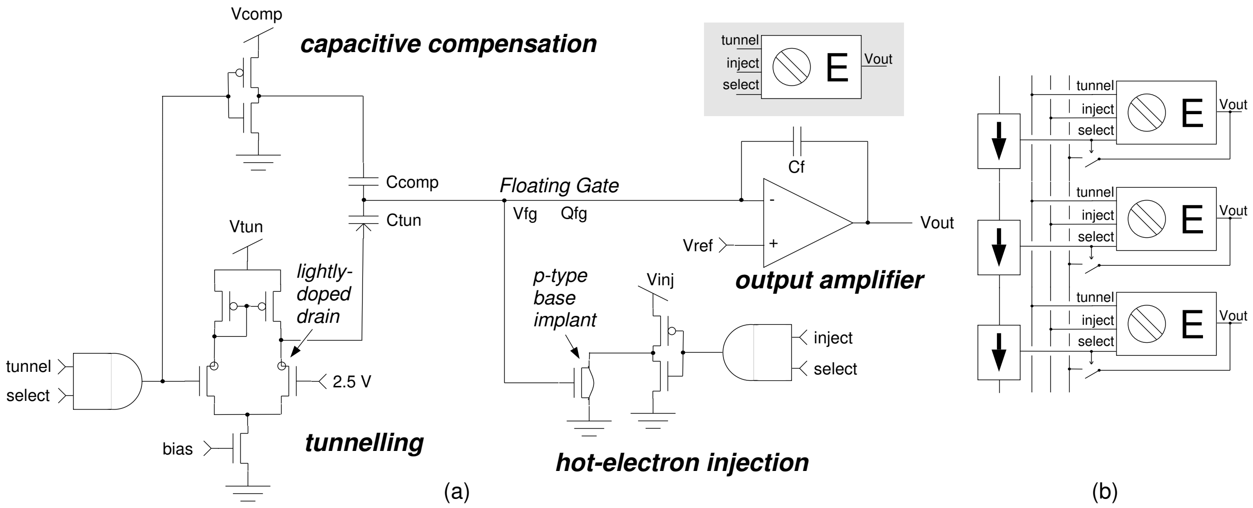

Figure 2. E-pot circuit schematic. (a) Single e-pot cell. The floating gate is connected to the negative input of an amplifier with feedback capacitor C_f. This pins the floating gate voltage to V_{ref}, and allows V_{out} to be moved from rail to rail by changing the charge of the floating gate Q_{fg}. Charge is removed from the floating gate through tunneling. The tunneling voltage is switched with a high-voltage differential amplifier built with weakly-doped-drain nFETs. The capacitive coupling of the tunneling voltage to the floating gate is counterbalanced by switching a lower voltage V_{comp} on a larger capacitor. Charge is added to the floating gate through hot-electron injection. We use p-type base implant to create an nFET with high threshold voltage, which is capable of injecting hot electrons onto the floating gate. (b) Array of e-pots with shift register used for addressing. The tunnel, inject, and select lines carry digital signals.

2. Small: We would like to get reasonable numbers of these devices on one chip in order to have many "virtual pins."

3. Few Pins: Because our goal is to save pins, a large array of e-pots should require few pins for I/O and biases.

4. Wide Voltage Range: Each e-pot should be able to store any voltage from ground to the power supply rail.

5. "Tweakable:" An e-pot's voltage should move up or down smoothly, just like turning the knob on a potentiometer.

6. Individually Addressable: We must have control over each e-pot, like an array of real potentiometers (*i.e.,* no global erase necessary).

7. High Programming Precision: More precision is always desirable.

Recent advances in floating-gate CMOS circuits open up the possibility for building on-chip non-volatile e-pots [1–6]. In this paper, we describe a floating-gate CMOS e-pot array that meets the requirements listed above.

2. CIRCUIT DESCRIPTION

The circuit schematic is shown in Figure 2. We designed and fabricated an array of e-pots in a commercially-available 2.0μm CMOS process. Configured in a 1-D array (see Figure 2b), each e-pot measures 69μm $\times$ 333μm including addressing circuitry. We constructed an array of 23 e-pots along the bottom of a 2.3mm $\times$ 2.3mm TinyChip. An e-pot array of arbitrary length consumes 11 pins for control signals and other biases. Once programming is completed, the array needs only three off-chip biases.

In order to make the e-pots non-volatile, we use floating-gate MOS transistors. A floating gate is a polysilicon node surrounded by SiO_2, trapping charge on the gate indefinitely. The floating gate is connected to the negative input of a high-gain amplifier with a feedback capacitor C_f. This effectively pins the floating gate voltage at V_{ref}, the positive input to the amplifier. By adding or removing charge from the floating gate, we can change the e-pot output voltage (see Figure 3). We have set e-pots to voltages from 10mV above ground to the power supply rail.

We can add electrons to the floating gate by the process of hot-electron injection. We use p-type base implant to create an nFET with a high threshold voltage (around 6V). We can operate this device in subthresh-

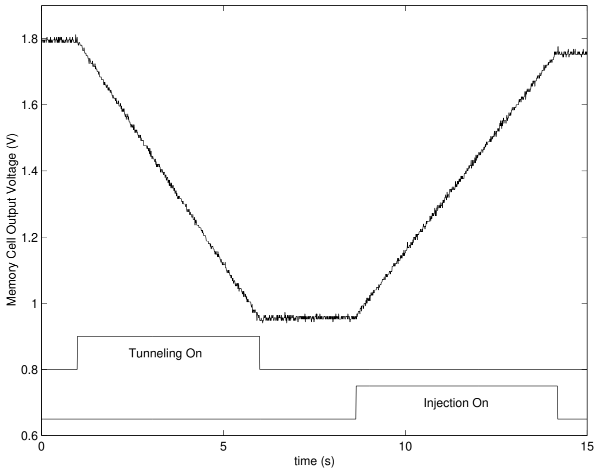

Figure 3. Experimental measurement illustrating e-pot operation. Digital signals control tunneling and hot-electron injection, moving the e-pot output voltage up or down smoothly.

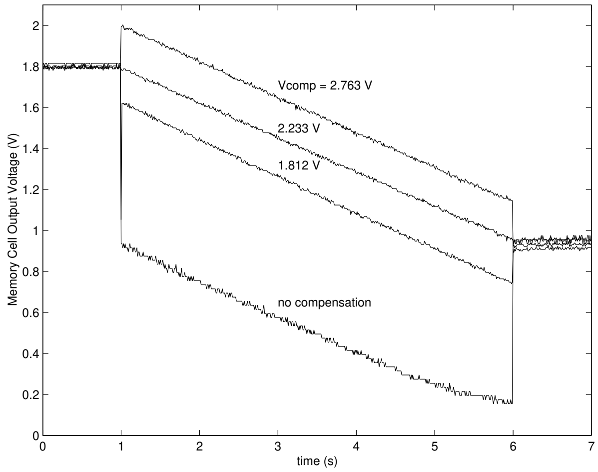

Figure 4. Illustration of compensating for capacitive coupling during tunneling. Normally, capacitive coupling through the tunneling junction produces large offsets in V_{out} during tunneling. By tuning V_{comp} to the correct value, we can compensate for this effect.

old with a gate voltage of 5.5V and inject electrons to the floating gate by raising the drain voltage to 4V. Electrons approaching the drain gather energy from the strong electric field, and some gain enough energy to surmount the 3.1eV energy barrier imposed by the gate oxide. These "hot" electrons are attracted to the floating gate, which is more positive than the drain. Because the floating gate must be around 5.5V, we use a 6V power supply for the e-pot and set the floating-gate voltage with the feedback circuit mentioned above. We could equally have used standard pFETs as injection elements [2].

We can remove electrons from the floating gate by the process of Fowler–Nordheim tunneling [7]. This process requires a high voltage (35–40V) source in order to create a sufficiently thin energy barrier that electrons can tunnel through the gate oxide. We switch high voltages on chip with a high-voltage differential amplifier built with lightly-doped-drain nFETs. We use well diffusion as the lightly-doped drain regions. The resulting high-voltage nFETs have breakdown voltages greater than 45V. In the CMOS process we use, it is not possible to create high-voltage pFETs, so we simply let the pFETs in the diffamp break down. This happens at a V_{ds} of 25V, so our high-voltage diffamp cannot completely "turn off". The output varies between around 40V in the "on" mode to around 15V in the "off" mode. However, the tunneling current depends exponentially on the reciprocal of the oxide voltage, so at $V_{tun} = 15$V, there is no observable tunneling current. It should be noted that after the e-pots are programmed, the tunneling voltage can be removed and there is no further need for a high-voltage source. By switching the tunneling and injection voltages locally, we achieve individually-programmable memory cells, with no need for a global erase. While programming individual e-pots, we measured no crosstalk to other e-pots due to tunneling or injection.

A major limitation of floating-gate memory cells that are programmed by tunneling results from parasitic capacitances coupling between the tunneling node and the floating gate. Although this capacitance is usually small compared to the total capacitance of the floating gate, the voltage change across this junction is large (typically 25–35V) when the tunneling voltage is switched on or off. This voltage change couples into the floating gate, causing a large voltage offset in the memory cell output for the duration of tunneling (see Figure 4). This makes for indirect and awkward programming.

In order to make our memory cells "tweakable," we compensate for this effect by swiching a lower voltage on a larger capacitor in the opposite direction. When we switch the tunneling junction to a high voltage, we simultaneously switch another node from a positive "compensation voltage" to ground. This node is coupled to the floating gate with a capacitance that is many times larger than the tunneling junction capacitance. The compensation voltage V_{comp} is set by an off-chip bias, so the capacitive coupling can be precisely nulled by the user (see Figure 4).

An earlier floating-gate memory cell was built by Diorio and colleagues [3]. Our e-pot design has four advantages over this design. (1) By using a high-gain differential amplifier in a feedback configuration we have explicit control of the floating gate voltage through V_{ref}. (2) We use an output amplifier with nFET inputs, so there is no leakage resulting from pFET hot-electron injection. (3) Our memory cells are configured as an addressable array with no crosstalk, and high-voltage switches in each e-pot make a global erase unnecessary. (4) The capacitive coupling inherent in tunneling is cancelled.

3. PROGRAMMING PRECISION

We used a comparator in a feedback loop with a memory cell in order to program a voltage precisely. Depending on the feedback connections, we could program a voltage by tunneling or by injection. The only limitation of this technique is the offset voltage of the comparator, typically less than 5mV for the particular circuit that we used.

In order to measure the precision of the memory cell, we programmed the same voltage 100 times and measured the error between the resulting voltage and the reference voltage. To quantify precision, we computed the standard deviation of the resulting voltages. The mean error voltage was around 5mV for the case of injection. Part of this could be caused by the comparator, but the rest is the result of capacitive coupling to the floating gate through the gate-to-drain overlap capacitance of the injection MOSFET. In principle, we can compensate for this offset in the same way that we deal with tunneling, but to do so, we would need to supply another off-chip bias. The mean error voltage in the tunneling case is dependent on V_{comp}, and can be nulled out. The histograms of the resulting error voltages (with the mean subtracted) are shown in Figure 5. The standard deviations of 450μV for injection and 470μV for tunneling correspond to greater than 13-bit precision over the 6V operating range of the e-pot. We made no attempts to measure this data in a low-noise environment, so this number is merely a lower bound on e-pot precision.

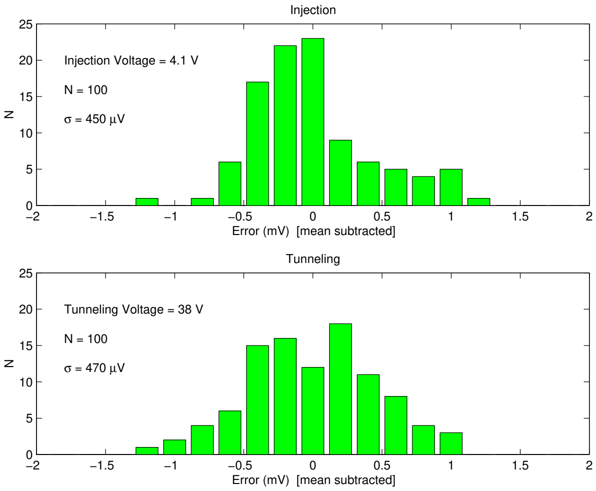

Figure 5. E-Pot precision. The e-pot was programmed to $V_{out} = 3$V using a comparator in a feedback loop. This was repeated 100 times for both tunneling and injection. (a) Histogram for programming with hot-electron injection. (b) Histogram for programming with tunneling. Both processes give a precision (6V range divided by σ) of greater than 13 bits.

ACKNOWLEDGEMENTS

This work was supported by the Center for Neuromorphic Systems Engineering as a part of the National Science Foundation's Engineering Research Center program. Reid Harrison is supported by an NDSEG fellowship from the Office of Naval Research.

REFERENCES

[1] P. Hasler, C. Diorio, B.A. Minch, and C. Mead, "Single transistor learning synapses with long term storage," *Proc. of the Intl. Symposium on Circuits and Systems*, Seattle, vol. 3, 1995, pp. 1660-1663.

[2] P. Hasler, B.A. Minch, C. Diorio, and C. Mead, "An autozeroing amplifier using *p*FET hot-electron injection," *Proc. of the Intl. Symposium on Circuits and Systems*, Atlanta, vol. 3, 1996, pp. 325-328.

[3] C. Diorio, S. Mahajan, P. Hasler, B.A. Minch, and C. Mead, "A high resolution non-volatile analog memory cell," *Proc. of the Intl. Conf. of Circuits and Systems*, Seattle, vol. 3, 1995, pp. 2233-2236.

[4] C. Diorio, P. Hasler, B.A. Minch, and C. Mead, "A complementary pair of four-terminal silicon synapses," *Analog Integrated Circuits and Signal Processing*, vol. 13, no. 1/2, 1997, pp. 153-166.

[5] 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.

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

[7] M. Lenzlinger and E.H. Snow, "Fowler–Nordheim tunneling into thermally grown SiO_2," *Journal of Applied Physics*, vol. 40, no. 1, 1969, pp.278–283.