

# The Floating Gate MOS Device as an Analog Element for Analog Trimming

Weinan Gao and W. Martin Snelgrove  
Dept. of Electrical Engineering  
University of Toronto, Canada, M5S 1A4  
(416) 978-3381 wgao@eecg.toronto.edu

**ABSTRACT** -This paper investigates the use of floating gate MOS devices for analog trimming. A floating gate MOSFET structure, in which tunneling occurs at the corners of a polysilicon slab, has been fabricated using a standard  $1.2\mu\text{m}$  n-well CMOS process. The analog storage performance of the floating gate is evaluated. A three-dimensional device simulator, DAVINCI, is used to characterize the field enhancement due to geometry. The charge-transfer behavior of floating gate MOS devices is thoroughly analyzed and modelled. This model describes the charging and discharging of the floating gate and provides insights to the design of analog floating gate devices. Feedback-based programming strategy is also discussed.

## I. INTRODUCTION

Floating gate MOS transistors currently have widespread applications as the storage elements in EPROM and EEPROM circuits. These devices are optimized for small area, high write/erase speed and low failure rate when used for digital memory [1] [2]. Due to their long-term charge retention, the floating gate devices would have potential applications in analog circuits [3-5] but they are not trivial to implement and need much fundamental work.

Our research is directed toward use of floating gate devices as compact and non-volatile elements for analog trimming. Among the desired features of such a device are bidirectional fine trimming, low programming voltages, good charge stability, and large trim range. Also, it is desirable to avoid the need for special processing to achieve ultrathin oxide film [6] or textured-surface polysilicon [7] since these processing features are expensive and often not available. Based on a previous attempt to build a tunneling injector in a  $2\mu\text{m}$  standard p-well CMOS process by L. R. Carley [3], we have designed and fabricated tunneling injectors which rely on the corners of a polysilicon slab and a floating gate MOS device to test the feasibility of floating gate analog trimming circuits. The implementation technology was CMOS4S, a  $1.2\mu\text{m}$  n-well standard CMOS process [8].

To analyze these tunneling injectors, three-dimensional field simulation is required. This study uses the three-dimensional device simulator DAVINCI [9] to analyze field enhancement effects induced by layout features.

To understand and control the transfer of charge onto and off the floating gate, it is necessary to develop a model which analytically describes charging and discharging of the floating gate via Fowler-Nordheim tunneling. Our model ties together most important parameters necessary for the analog floating gate MOS device design such as tunneling oxide thickness, tunneling area, critical coupling capacitances, programming time, programming voltages and initial floating gate voltage.

## II. TUNNELING INJECTOR STRUCTURES AND CHARACTERIZATION

The floating gate MOS device in our study consists of a tunneling injector, a sense transistor and a polysilicon bootstrap capacitor (Figure 1). The tunneling injector enables the flow of electrons onto or off the floating gate by means of Fowler-Nordheim tunneling, and thereby sets the analog voltage on the floating gate which controls the operation of the sense transistor.

Normally, special processing requirements such as ultrathin tunneling oxide or textured-surface polysilicon are required to induce Fowler-Nordheim tunneling at a reasonable voltage ( $<20\text{V}$ ) [6]. It has been suggested that corners introduced by lithographic features in the polysilicon layer can enhance the electric field in the oxide film [3]. To examine the influence of corners on the tunneling characteristic, a test chip which consisted of several tunneling injectors with different numbers of corners was designed and fabricated in CMOS4S. This process has a  $250\text{\AA}$  gate oxide thickness and a poly-poly oxide thickness of  $400\text{\AA}$ . The layouts of these tunneling injectors are shown in

Figure 2, in which the polysilicon slab in each injector ends in the middle of a diffusion region. According to electromagnetic field theory, the electric field is enhanced at the comers of this polysilicon slab [10].

The tunneling current is theoretically characterized by the Fowler-Nordheim equation[ 11]

$$J_{tun} = \alpha E_{tun}^2 \exp\left(-\frac{\beta}{E_{tun}}\right) \quad (1)$$

where  $J_{tun}$  is the current density, and  $\alpha$  and  $\beta$  are constants determined from experimental data.

Figure 3 shows a typical measured tunneling current characteristics for electrons being injected onto (forward) or removed (reverse) from the floating gate as a function of the voltage difference between the floating gate and the diffusion region. It can be seen that the voltages required to induce the tunneling current in the forward and reverse directions are about 12V and 15V, respectively. They are much lower than the voltages for inducing tunneling in a structure without **comers**<sup>1</sup>. We didn't find tunneling in the comerless structure up to 25V. The tunneling current curve is asymmetric with slightly higher reverse tunneling voltage. To better characterize the tunneling performance of tunneling injectors, these measured data are replotted on Fowler-Nordheim axes ( $\log(I/V^2)$  vs.  $(V^{-1})$ ) in Figure 4. The curves fit the Fowler-Nordheim equation well since the experimental data lie close to a straight line

We suspect that the asymmetry of tunneling is due to reverse behavior being defined by the comers of the polysilicon slab itself, while forward tunneling is defined by comers in the underlying diffusion.

From the forward Fowler-Nordheim plot,  $\alpha$  and  $\beta$  can be extracted as  $1.25 \times 10^{-6} \text{ A/V}^2$  and  $2.67 \times 10^8 \text{ V/cm}$ , respectively. In the reverse direction,  $\alpha$  and  $\beta$  are  $1.88 \times 10^{-6} \text{ A/V}^2$  and  $3.15 \times 10^8 \text{ V/cm}$ , respectively.

Without field enhancement, the tunneling-oxide electric field is given by

$$E_{tun} = \frac{|V_{tun}|}{X} \quad (2a)$$

where  $V_{tun}$  is the voltage drop across the oxide and X is its thickness. Field enhancement increases electric fields near comers by a constant factor, so

$$E_{tun} = \kappa \frac{|V_{tun}|}{X} \quad (2b)$$

where  $\kappa$  expresses the degree of field enhancement. For the experimental tunneling injectors, the reduction in tunneling from the "comerless" case indicates that  $\kappa$  in the forward and the backward tunneling direction is about 2.0 and 1.6, respectively.

Our experimental data showed a great deal of variability in I-V characteristics between structures. Structure (d) and (e) have better uniformity over measurements of five different dice. This may be due to having more comers in Structure (d) and better control of comer radius in fabricating structure (e). Another important phenomenon is that measured tunneling current is not proportional to the number of comers. Both of these results indicate either the exact shape of the three-dimensional comers or the extent of interface trap states being very important in determining the degree of field enhancement. Therefore one of disadvantages of this tunneling injector is that it is sensitive to process variations because the process features strongly affect the exact shape of these 3-D comers. We don't know how consistent parameters are from run to run, but have the advantage of using a single foundry.

It has been recently reported that tunneling can also occur through the oxide between the two polysilicon layers normally used for capacitors and that tunneling voltage is more reliable and controllable if there is a large tunneling

---

<sup>1</sup> For the structure without any comers in our test chip, no tunneling current was observed even up to 25 volts.

area [12]. The poly1-poly2 tunneling injector may be useful as a substitute for the poly-diffusion tunneling injector in our floating gate analog MOS device structure.

Because it is the comers of the polysilicon that create enough field enhancement to dramatically lower the tunneling voltage, a three-dimensional field analysis is required to predict the degree of field enhancement and other three-dimensional effects. Figure 5 shows electric field distributions simulated with DAVINCI in the oxide around the rectangular comers on applying 14 volts to the polysilicon electrode. It can be seen that the electric fields around the comer are stronger than the electric fields far away from the comer, and stronger than those without comers. The degree of field enhancement predicated by this simulation is about 1.5, lower than the data given in[3]. The reason may be the description of the exact shape of three-dimensional comers, which is hard to know due to variability in fabrication. Further, DAVINCI cannot analyze arbitrary-shaped comers. To better predict tunneling performance and the impact of the shape of the three-dimensional comers on tunneling, an advanced three-dimensional device simulator which supports arbitrary grid structure and tunneling characteristics is required.

### III. FLOATING GATE ANALOG STORAGE

The objective of floating gate analog trimming is to establish precise voltage levels on the gate of the sense transistor, which is dependent on the amount of stored charge on the floating gate. The floating gate is completely isolated by surrounding layers of oxide and acts as a capacitor with long-term charge retention capability [2], [3]. For programming the floating gate MOS device, a sufficiently high voltage across the tunneling injector causes a tunneling current to flow through the tunnel oxide, into or out of the floating gate, which changes the amount of charge stored on it.

Essentially, the two variables that may be used to vary the amount of charge and hence the voltage stored on the floating gate are: (1) programming voltage, (2) programming time. The programming voltage controls the target values on the floating gate and programming time controls how close to this target value the programming comes.

Floating gate analog storage is designed to handle trim voltages that require high precision and good stability. Therefore, the variability of their behavior is a major concern. It is well known that most floating gate digital memories have data charge retention exceeding 10 years at room temperature [1]. Since the floating gate MOS device here employs a standard gate oxide as the tunnel oxide, its retention should be very good. It has been experimentally demonstrated that the charge loss of these floating gate memories using a standard gate oxide for tunneling is less than 0.1 percent in ten years at operating and storage temperatures up to 100°C [3].

### V. MODELLING OF CHARGE-TRANSFER CHARACTERISTICS

In order to gain insight into the basic charge-transfer characteristic of floating gate MOS devices, a capacitive equivalent circuit shown in Figure 6, is developed. From this equivalent circuit, the floating gate voltage,  $V_{FG}$ , can be formulated using superposition as

$$V_{FG} = P_{CG}V_{CG} + P_I V_{INJ} + P_S V_S + P_D V_D + P_B V_B \pm \frac{Q_{FG}}{C_{FG}} \quad (3)$$

with

$$C_{FG} = C_G + C_{FI} + C_{FLD} + C_S + C_D + C_C \quad (4a)$$

$$P_{CG} = \frac{C_G}{C_{FG}}, P_I = \frac{C_{FI}}{C_{FG}}, P_S = \frac{C_S}{C_{FG}}, P_D = \frac{C_D}{C_{FG}}, P_B = \frac{C_{FLD} + C_C}{C_{FG}} \quad (4b)$$

where  $P_{CG}, P_I, P_S, P_D$  and  $P_B$  are the coupling ratios from the control gate, the current injector drain, the source of the n-channel transistor, the drain of the n-channel transistor and the substrate to the floating gate, respectively.  $Q_{FG}$  is the total charge on the floating gate, with "+" used for positive charge and "-" for negative.

Based upon the above equivalent circuit and the Fowler-Nordheim tunneling equation, analytical expressions can be derived to show the change in floating gate voltage as a function of programming voltage and time in each charging (injecting electrons onto the floating gate) or discharging (removing electrons from the floating gate) operation [13].

The change in floating gate voltage during charging is

$$\Delta V_{FG}(t) = -V_{FG}(0) + V_{INJ}(t) + \frac{M}{\ln[\exp(\frac{M}{V_{FG}(0) - V_{INJ}(0)}) + \beta NPt]} \quad (5a)$$

and the change in floating gate voltage during discharging is

$$\Delta V_{FG}(t) = -V_{FG}(0) + V_{INJ}(t) - \frac{M}{\ln[\exp(-\frac{M}{V_{FG}(0) - V_{INJ}(0)}) + \beta NPt]} \quad (5b)$$

where

$$V_{FG}(0) = P_{CG}V_{CG} + P_I V_{INJ}(0) + P_S V_S + P_D V_D + P_B V_B \pm \frac{Q_{FG}(0)}{C_{FG}} \quad (6)$$

$$\beta X_{tun} = M, \quad \frac{\alpha A_{tun}}{C_{FG} X_{tun}} = N, \quad P = \frac{C_{FG} + C_{INJ}}{C_{FG} + C_{FI}} \quad (7)$$

Eqs. (5a) and (5b) characterize the net electron accumulation and the net electron loss, respectively. If the voltage at the injection electrode,  $V_{INJ}$ , is a constant, the above equations reduce to

$$\Delta V_{FG}(t) = -V_{FG}(0) + V_{INJ} + \frac{M}{\ln[\exp(\frac{M}{V_{FG}(0) - V_{INJ}}) + \beta Nt]} \quad (8a)$$

for the charging process, and

$$\Delta V_{FG}(t) = -V_{FG}(0) + V_{INJ} - \frac{M}{\ln[\exp(-\frac{M}{V_{FG}(0) - V_{INJ}}) + \beta Nt]} \quad (8b)$$

for the discharging process.

Furthermore in the field enhancement situation, the equations (15a) and (15b) are simply modified as

$$\Delta V_{FG}(t) = -V_{FG}(0) + V_{INJ}(t) + \frac{M}{\kappa_1 \ln[\exp(\frac{M}{\kappa_1 [V_{FG}(0) - V_{INJ}(0)]}) + \beta NPt]} \quad (9a)$$

for net electron accumulation, and

$$\Delta V_{FG}(t) = -V_{FG}(0) + V_{INJ}(t) - \frac{M}{\kappa_2 \ln[\exp(-\frac{M}{\kappa_2 [V_{FG}(0) - V_{INJ}(0)]}) + \beta NPt]} \quad (9b)$$

for net electron loss.  $\kappa_1$  and  $\kappa_2$  characterize the degree of field enhancement in forward and reverse directions, respectively.

The above equations, giving  $V_{FG}(t)$  versus time, are related to all of the most important parameters. Therefore, they are useful for the design and analysis of floating gate analog MOS devices. They can also be used to evaluate the charge-transfer characteristics and the controllability of analog floating gate devices. As an example of the application of this model, the programming rate of the floating gate in charging is simulated, with results shown in Figure 7. Data are shown with programming voltages of 15V, 16.5V and 17V applied to the control gate. The parameters we chose were based on the layout parasitics of our fabricated floating gate analog MOS devices and the measured characteristics of our tunneling injectors; they are:  $\alpha=1.25\times 10^{-6} \text{ A/V}^2$ ,  $\beta=2.57\times 10^8 \text{ V/cm}$ ,  $\kappa=2$ ,  $X_{tun}=250\text{\AA}$ ,  $A_{tun}=1\mu\text{m}^2$ ,  $C_{FG}=1\text{pF}$ ,  $C_{INJ}=5\text{fF}$ ,  $V_{FE}=12\text{V}$ ,  $P_{CG}=0.8$ . As seen from this figure, the charging process is very fast; the changes in floating gate voltage achieved in  $1\mu\text{s}$  are -0.054V, -0.361V and -0.834V with the three control voltages. We note that the change in floating gate voltage increases monotonically with increasing charging time. Also, from this simulation, we can see a basic trade-offs: higher programming levels result in a larger trim range (the extent of target values that may be achieved), but make it harder to control trim step size.

## V. FEEDBACK-BASED PROGRAMMING

The operation of the floating gate device needs control circuitry to provide both positive and negative programming voltages to inject charge onto or remove it from the floating gate. An efficient programming strategy is to embed the floating gate MOS device into a feedback loop[14]. The local feedback circuitry measures the extent of the uncompensated offset in the system and then decides **whether** the trimming is needed, as well as which type of programming pulses is required.

A conceptual diagram of our feedback-based programming strategy is shown in Figure 8. Programming pulses are generated by high voltage driving circuitry, and the UP/DOWN counter decides programming direction. Accordingly, the amount of charge on the floating gate is increased or decreased to achieve a target value which is used to cancel the offset of the uncompensated circuit.

## VI. CONCLUSION

The work presented here indicates that floating gate devices are very usable as analog trimming elements. The successful implementation of tunneling injectors with a number of comers demonstrates that floating gate trimming methods can be implemented in a standard CMOS IC process, although more work is still needed to provide a more reliable and controllable floating gate device. The charge-transfer model provides a powerful tool for the design and analysis of analog floating gate devices.

## ACKNOWLEDGEMENTS

This work was supported by MICRONET and NSERC Canada.

## REFERENCES

- [1] P. Prince, *Semiconductor memories*, England, Addison Wisley, 1992.
- [2] H. Maes, J. Witters and G. Groeseneneken, "Trends in non-volatile memory devices and technologies," in *Proc. 17th Euro. Solid State Devices Res. Conf.* pp. 743-754, 1987.
- [3] L. Richard Carley, "Trimming analog circuits using floating-gate analog MOS memory," *IEEE J. Solid-State Circuits*, vol. SC-24, pp. 1569-1575, Dec. 1989.
- [4] Säckinger and W. Guggenbühl, "An analog trimming circuit based on a floating-gate device," *IEEE J. Solid-State Circuits*, vol. SC-23, pp. 1437-1440, Dec. 1988.
- [5] B. W. Lee, B. J. Sheu and H. Yang, "Analog floating-gate synapses for general-purpose VLSI neural computation," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 654-658, June 1991.
- [6] W. S. Johnson, G. L. Kuhn, A. L. Rinninger and G. Perlegos, "16-K EE-PROM relies on tunneling for byte-erasable program storage," *Electronics* vol. 53, pp. 113-117, Feb. 1980.

- [7] J. Drori et al, "A single 5-volt supply nonvolatile static RAM," in Dig. Tech. Papers *IEEE Int. Solid-State Circuit Conf.*, pp. 7148-149, Feb. 1981.
- [8] D. Brown and A. Scott, "Design rules and process parameters for the Northern Telecom CMOS4S process," CMC, Kingston, Canada, 1990.
- [9] TMA Inc., "DAVINCI Three-dimensional device analysis program," 1991.
- [10] H. A. Haus and J. R. Melcher, *Electromagnetic Fields and Energy*, Englewood Cliffs, NJ Prentice Hall, pp. 177-178, 1989.
- [11] M. Lenzlinger and E. H. Snow, "Fowler-Nordheim tunneling into thermally grown SiO<sub>2</sub>," *J. Appl. Phys.*, vol. 40, pp. 278-283, Jan. 1969.
- [12] D. A. Durfee and F. S. Shoucair, "Comparison of floating gate neural network memory cells in standard VLSI CMOS technology," *IEEE Trans. Neural Networks*, vol. 3, pp. 347-353, May 1992.
- [13] W. Gao, "Analog trimming using floating gate devices," M.A.Sc Thesis, University of Toronto, Toronto, Canada 1992.
- [14] A. Shoval, D. A. Johns and W. M. Snelgrove, "Median-based offset cancellation circuit technique," *1992 IEEE INT. Symp. Circuits Syst., San-Diego, California*, vol. 4, pp. 2033-2036, May, 1992.

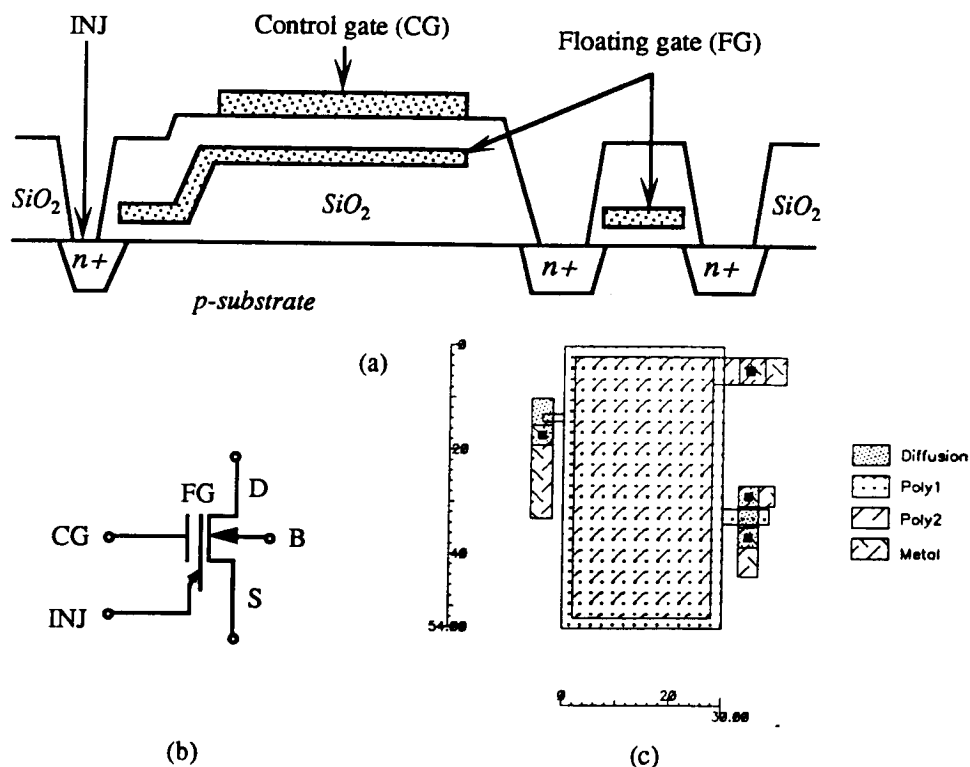


Figure 1. Analog floating gate MOS device structure: (a) cross section; (b) circuit symbol; (c) layout diagram.

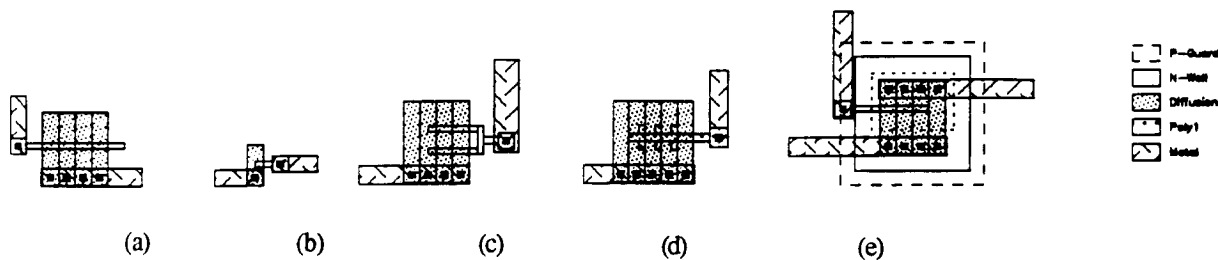


Figure 2. Layouts of tunneling injectors: (a) no corners; (b) two corners; (c) four corners; (d) fourteen corners; (e). two corners over a p<sup>+</sup> diffusion region.

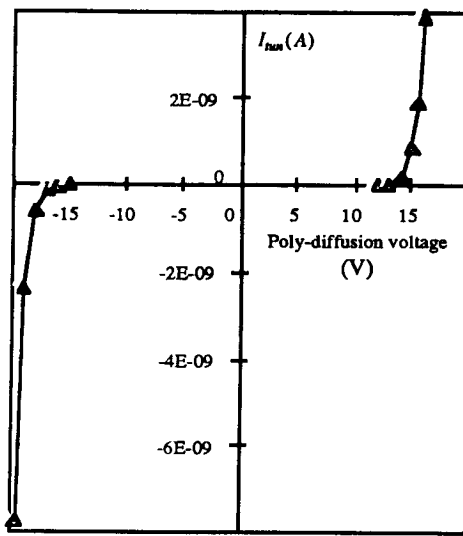


Figure 3. The measured I-V characteristic of a tunneling injector.

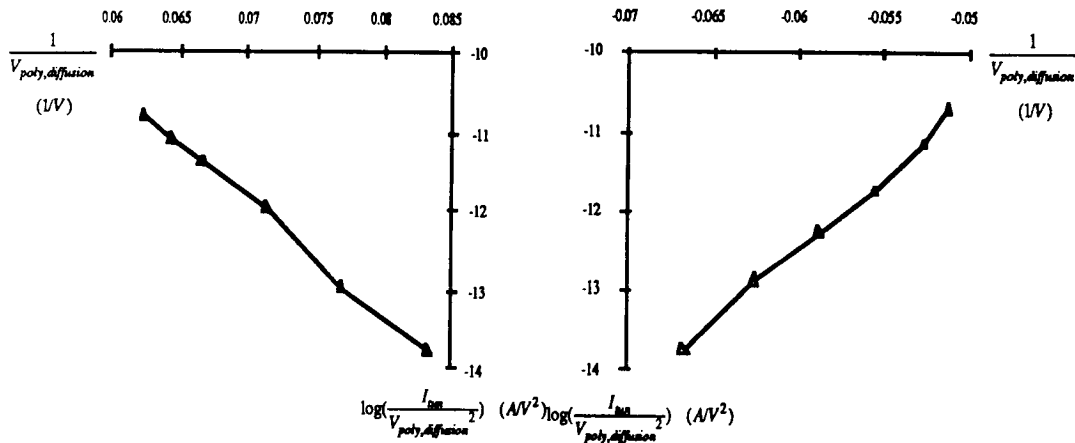


Figure 4. Fowler-Nordheim plot of the Figure 3: (a) forward direction; (b) reverse direction.

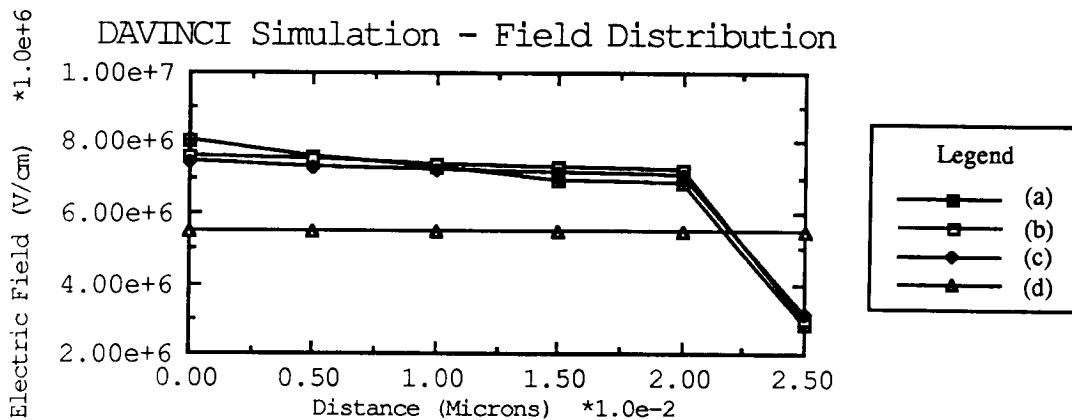


Figure 5. 3-D simulation for field distribution in oxide in case of a rectangular polysilicon corner for the characterization of field enhancement: (a) at the corner; (b)  $0.05\mu\text{m}$  away from the corner along the poly slab; (c)  $0.1\mu\text{m}$  away from the corner along the poly slab; (d)  $0.2\mu\text{m}$ ,  $0.3\mu\text{m}$  and  $0.4\mu\text{m}$  away from the corner along the poly slab, and the field distribution under the poly slab without corners.

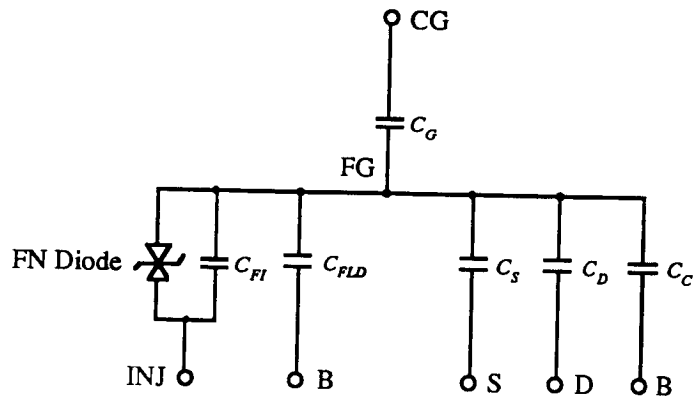


Figure 6. A simplified capacitive equivalent circuit of the floating gate MOS device.  $C_G$  is the floating gate to the control gate capacitance,  $C_{FI}$  is the overlap capacitance from the floating gate to the injection electrode of the tunneling injector,  $C_{FLD}$  is the field oxide capacitance from the floating gate to the substrate,  $C_S$  is the overlap capacitance from the floating gate to the source,  $C_D$  is the overlap capacitance from the floating gate to the drain and  $C_C$  is the capacitance from the floating gate to the channel.

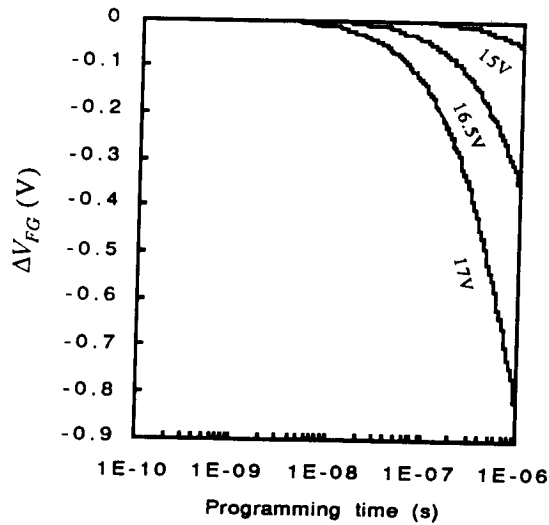


Figure 7. Simulation for the programming rate of charging operation with control voltages of 15V, 16.5V and 17V.

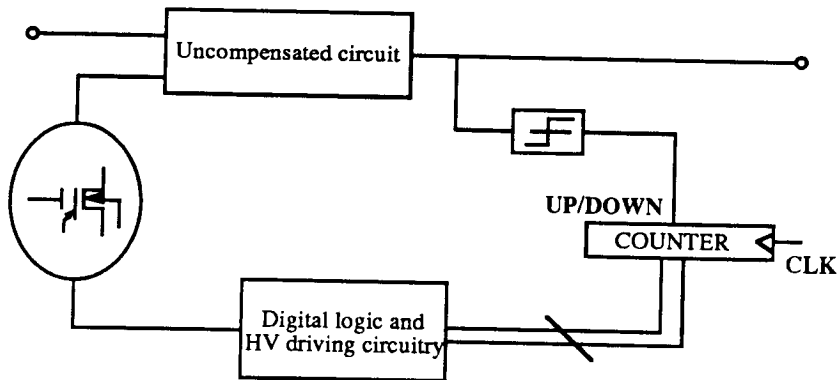


Figure 8. A conceptual diagram of a feedback-based programming strategy.