Trade-offs for High Yield in 90 nm Subthreshold Floating-gate Circuits by Monte Carlo Simulations

Jon Alfredsson
Department of Information Technology and Media,
Mid Sweden University
SE-851 70 Sundsvall, Sweden
e-mail: jon.alfredsson@miun.se

Snorre Aunet
Department of Informatics, University of Oslo
Postbox 1080 Blindern,
0316 Oslo, Norway
e-mail: sa@ifi.uio.no

***Abstract* - The work described in this paper is performed to estimate the influence of statistical process variations and transistor mismatch that occurs in fabrication and affect floating-gate digital circuits. These effects will affect and reduce "yield" (percentage of fully functional circuits). Monte Carlo simulations have been performed in a 90 nm to estimate the yield for manufactured floating-gate circuits running with subthreshold power supply. The power supply, floating-gate charge voltage (V_{FGP} and V_{FGN}) and transistor sizes have been varied during the simulations and the yield has been observed. The simulation results shows that by doubling the minimum size transistors (length and width) the yield can be much better than for minimum size version. A yield of 100% can though not be expected if the power supply is scaled down below 250 mV.**

I. INTRODUCTION

Subthreshold operation, where the power supply, $V_{dd,}$ is below threshold voltage, V_{th}, is known to be more effective in terms of lower power consumptions than other types of low-power techniques [1]. Subthreshold have also been proposed for signal processing floating-gate circuits [2] and implemented with threshold elements to get improved matching [3]. For an overview of different floating-gate techniques, see [4].

To reduce power consumption and increase manufacturability, including design of circuits that sustain process parameter variations, are two of the most important "grand challenges" for future nanotechnologies that have been determined for the semiconductor industry by ITRS [5]. A reduced power consumption will not only extend the operational time for battery driven applications but it will also for example make it possible to reduce costs related to heat removal and cooling equipment. Getting better control of process variations and transistor mismatches can improve the percentage of working circuits from fabrication which will lead to reduced manufacturing costs and a higher "yield" (% of the produced circuits that works as specified).

Floating-gate MOS (FCMOS) has previously been proposed as a techniques to reduce power consumption for signal processing circuits [2],[6],[4]. Other proposals for reduced power consumption suggest reducing the power supply to subthreshold region [7],[8],[9] and by combining these methods, it has also been suggested that FGMOS can be a good alternative for use in subthreshold [2],[10].

When CMOS technologies have evolved with smaller gate lengths, reduced gate-oxide thickness has also led to increased current leakage through the transistor's gate. Floating-gate circuits are therefore preferred to have sufficiently thick gate-oxide, which does not match well with using newer process technologies with thin gate-oxide [11]. Larger gate leakage currents will affect the performance of FGMOS circuits in many ways such as speed, power consumption and functionality. Previous work in this area have also shown that FGMOS with a higher fan-in than three will not have better performance than CMOS [12] and in [13] it is shown that FGMOS full-adder circuits of this type can have an advantage compared to CMOS in terms of Energy-Delay Product (EDP) performance with a properly made design.

The aim of the work has been to estimate and summarize how statistical process variations and transistor mismatches will affect the performance of floating-gate circuits working with subthreshold power supply voltages. Especially, they will affect the signal voltage levels and since reducing V_{dd} is the single most powerful thing to do in order to reduce power consumption [14], we wanted to find out what is minimum V_{dd} when considering these process variations [15].

In this paper we have been looking at how trade-offs between V_{dd}, floating-gate voltages and transistor sizes can be done to increase yield. We have performed Monte Carlo simulations of a full-adder that use building blocks with a maximum fan-in of three [13] and comparison with an mirrored-gate implementation. Two different transistor dimensions have been used for the floating-gate implementation and our results indicate that minimum size transistors can not be used in modern nanoscale CMOS technologies. The trade-offs we have found from the simulations can be used to design for increased manufacturability and as far as we know, this is the first publication on floating-gate circuits exploiting Monte Carlo simulations.

II. FGMOS CIRCUITS

FGMOS is a technique to shift threshold voltage (V_{th}) seen from the driving node on the transistors. By shifting V_{th}, FGMOS circuits can be modified for operation in a wide range of areas and applications [16]. Speed can be

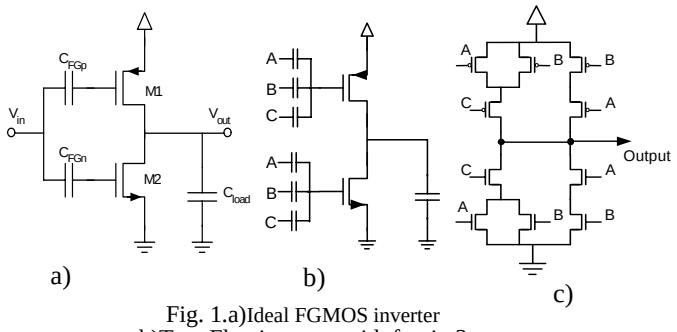

Fig. 1.a)Ideal FGMOS inverter
b)True Floating-gate with fan-in 3
c)Mirrored Gate element with fan-in 3

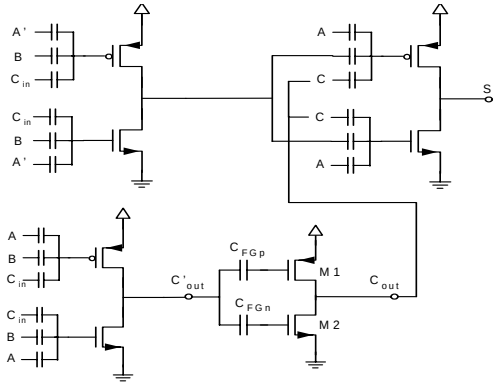

Fig. 2 Full-Adder design with the floating-gate elements.

traded for power so that different power consumption, PDP (Power-Delay Product) and EDP is achievable [4],[17].

To control this FG-charge, several methods have been developed. One is hot-electron injection [4],[16]. Other methods for control have been based on eraseable [6], UV-reconfigurable [2] or Fowler-Nordheim (FN) tunnelling [18]. FN-tunnelling can also combine with Hot-electron injection [19].

The FGMOS technique use normal MOSFET transistors and can be fabricated in a normal CMOS process. If the gate is completely isolated, it is called true Floating-gate and if it have some leakage currents its said to be pseudo- or quasi floating-gate. True floating-gate circuits are manufactured for long-term storage of information since the floating-gate charges are intended to be retained permanently or until reprogramming. Figure 1a shows a true (ideal) FGMOS inverter designed with two floating-gate transistors

Depending on the value of the floating-gate voltage (V_{FG}), the effective threshold voltage for the circuit will vary. One of the easiest ways to determine proper V_{FG}s via simulations is to use a specially designed balancing scheme. This will determine how V_{FGn} for the nMOS should be selected for each value of V_{FGp} for pMOS [20]. When a V_{FG} pair has been selected it is normally programmed via some of the methods mentioned above and then kept fixed during operation.

With today's progress in technology scaling it will be increasingly difficult to design functional true floating-gate circuits with the most recent processes than ever before. This is mostly due to decreasing gate-oxide thickness and increasing leakage currents in the transistors. Thicker gate-oxide is therefore needed for true floating-gate designs because some charge

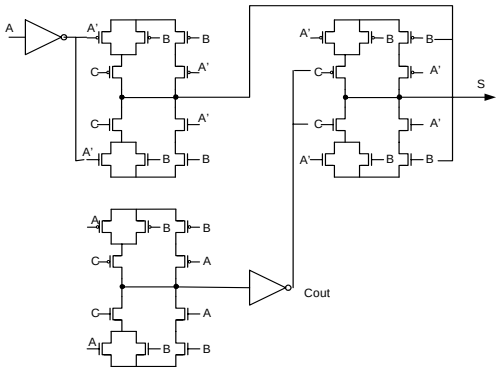

Fig. 3. Mirrored gate Full Adder topology.

leakages from the gate node are always expected to occur.

Floating-gate circuits have the advantage compared to other digital design techniques that several of the basic logic functions (NAND, NOR, NOT) and a few other commonly used gates (i.e. CARRY') can be designed with only two transistors per gate with the help of floating-gate capacitances representing the fan-in number [3]. According to previous research the fan-in can not be higher than 5 to be functional (at 250 mV V_{dd}) [12]. Even further restrictions are applied if the EDP performance should be better than for CMOS. The fan-in is then restricted to a maximum of 3.

III.SIMULATIONS

Figure 1b shows a true floating-gate element with fan-in 3. It is a minority-3 gate. This element has been used as a basic building block in the full-adder we have been simulating. For comparison reasons we have also been simulating a mirrored gate CMOS full-adder based on threshold elements seen in figure 1c and as stated before, it is a circuit design with better performance in terms of switching speed and lower power consumption (shorter gate delay and better EDP) than performance of similar standard static CMOS circuits [21].

The topology of the full-adder design used in our simulations can be seen in figure 2. The same topology has been used for both the FGMOS and mirrored gate full-adder and the schematic for the mirrored gate version can be seen in figure 3.

The simulations in this work are basically Monte Carlo simulations that we have performed in Cadence using the Spectre simulator and a 90 nm CMOS process technology from CMP [22]. The reason for this choice of process is because the previously used 130 nm process has no statistical data to use in simulation of process variations. The transistors have in all simulations been of high-V_t type, (hvt, low-leakage), and the transistor sizes have been chosen either to a minimum or double minimum possible length and width such as the circuits is balanced and an input signal of $V_{dd}/2$ gives an output signal of $V_{dd}/2$. The Monte Carlo simulations have been simlated with 100 runs. It will be enough to obtain a relatively good estimation accuracy since no absolute results are required.[25]

The Monte Carlo simulations take the statistical process variation and transistor mismatch into account when it simulate a circuit [26],[27]. These statistical variations will affect the

signal levels and in order for a circuit to be fully functional it has to meet requirements for the signal's voltage levels [15].

The input and output signal noise margin requirements for zero and one have in our simulations been set according to 25% and 75% of signal swing respectively [3]. Reference [26] shows an illustration of how this is done. That is, for example the Input High Voltage (V_{IH}) and Output High Voltage (V_{OH}) signals must be larger than 75% of V_{DD} to be considered as a One and accordingly, the Input Low Voltage (V_{IL}) and Output Low Voltage (V_{OL}) must be lower than 25% of V_{dd} in order to be considered as a Zero. Any failure to meet these requirements of zeroes and ones during the Monte Carlo simulations will be considered as a faulty circuit and affect the yield. The yield is, in our simulations, defined as the percentage number of circuits that have full functionality for all inputs. A circuit is considered fully functional when all output signal levels meets the requirement for a specific sequence given input signals. The Yield estimated from our simulations will not provide any absolute values but it tells what voltage limit it is to possible not reduce below to recieve a specific yield.

In the Monte Carlo simulations we have performed, 100 runs have been simulated for each V_{dd} and V_{FG}-pair (combination of nMOS and pMOS floating-gate charge voltage). The input signal stimuli sequence to the full-adder has been a set of binary counting bits A, B and C_{in} from 0,0,0, to 1,1,1.

IV.RESULTS

Figure 4 and figure 5 show the results of the Monte Carlo simulations. It shows the yield for different V_{dd} and different floating gate voltages (V_{FGp}). Figure 4 is the simulations with minimum transistor sizes (for a balanced input/output of $V_{dd}/2$) and figure 5 shows the simulation results with the size doubled.

Figure 4 shows that 100% yield is not achievable for FGMOS full-adders with minimum sized transistors but more than 80% yield can be achieved for certain V_{FGp} when supply voltage is scaled down to 350 mV.

If we instead double the transistor's lengths and widths, like in Figure 5, we can scale down to 250 mV and achieve 100% yield for certain choices of V_{FGp}. To scale down V_{dd} below 250 mV for FGMOS seems to be a bad idea since the yield then drops dramatically for both sizes of the full-adder.

Monte Carlo simulation results for the CMOS Mirrored-Gate full-adder can be seen in table 1. It shows the yield in percentage of functional circuits for different V_{dd} and with the same transistor sizes as in the case with the FGMOS full-adder. The cicuit is not expected to have 100% yield for V_{dd} under 150 mV (min. size) and under 100 mV (double size version).

V. DISCUSSION

In order to find out how performance for subthreshold FGMOS circuits will be when manufacturing problems like statistical process variations and transistor mismatches are taken into account we have been simulating full-adder circuits of different sizes and with different supply voltage.

For each current level that can flow through an FGMOS circuit there is only one corresponding combination of floating-gate voltages for the nMOS and pMOS transistors.

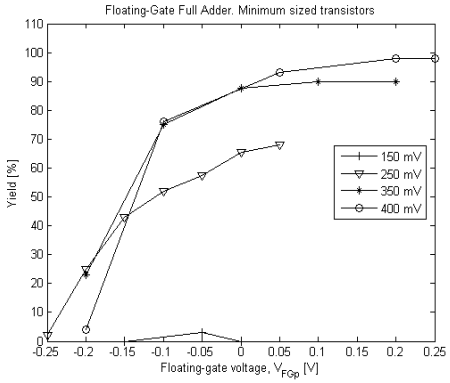

Fig. 4. Yield for a floating-gate full-adder working in subtrehshold with different Vdd and VFGp. Minimum Transistor sizes have been used.

The simulations show that the if the nMOS and pMOS transistors are chosen to minimum sizes it is not possible to have the circuits with 100% yield with subthreshold V_{dd}.

Table 1. Yield for Mirrored-Gate Full-Adder.

Vdd	100 mV	150 mV	250 mV	350 mV
Min Size	0%	99%	100%	100%
Double Size	100%	100%	100%	100%

Minimum transistor sizes for FGMOS have been proposed in previous work, [23], but as figure 6 shows it might be better to choose Mirrored-gate CMOS when minimum size transistors should be used in order to get higher yield. With upscaling the transistors by a factor 2, higher yield is also expected according to our simulations but even then, the simulation results from FGMOS estimate that the V_{dd} can not be reduced further below 250 mV and still achieve 100% yield. Looking at figure 4 and figure 5 we can also see how different V_{FG}-pairs and V_{dd} can give a wide variation of yield for FGMOS.

To notice is also that mirrored gate always show higher or equal yield in our simulations. Even for 2 times minimum sized transistors FGMOS can not achieve 100% yield for V_{dd} below 250 mV and at those levels some other technique should be the choice. Reference [27] gives an example of a system wih such a technique that was previously manufactured.

It should also be noticed that the estimated yields from our simulations are only valid for small simple circuits. When the ciruicts are connected to bigger systems on a chip, the yield will be affected significantly and reduced.

VI.CONCLUSION

With minimum size full-adder design, our Monte Carlo simulations show that a proper choice of floating-gate charge can give yield between 90-100% for power supplies of 350 mV and 400 mV. With 250 mV V_{dd}, a max. yield of 70% is achievable.

If the power supply should be scaled down below 250 mV FGMOS can not be recommended for use since in all of the simulations had problems with functionality.

However, if the transistor sizes are scaled up with a factor of

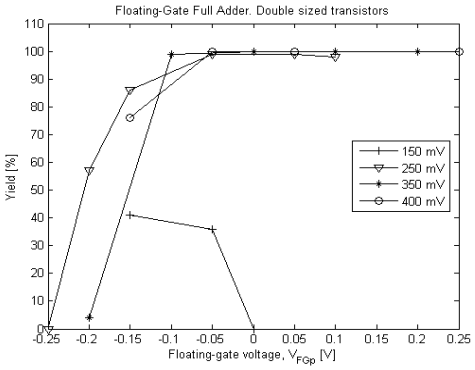

Fig. 5. Yield for a floating-gate full-adder working in subtrehshold with different V_{dd} and V_{FGp}. The transistor lengths and widths have been double minimum size design.

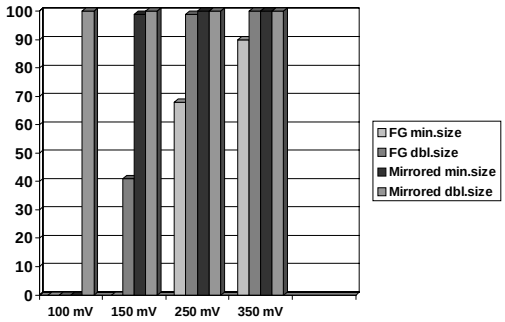

Fig. 6. Bargraph showing Yield for FGMOS and Mirrored gate with minimum and double minimum sized transistors for different power supplies.

2, there is a large improvement for the yield. Still there seems to be problem with functionality when V_{dd} is reduced below 250 mV. A yield of 100% can though not be expected to be achieved for power supply scaled down below that value.

When it comes to the simulated mirrored-gate CMOS full-adder it have better or equal yield than FGMOS for all power supplies. For the smallest V_{dd} (below 250 mV) mirrored gate seems to be a better choice than FGMOS if only the yield is to be considered. The simulation shows that it is possible to make it fully functional for certain cases when the power supply scales down to 100 mV and the transistor sizes are doubled. This is also a limit in power supply scaling according to [24].

VII.REFERENCES

[1] H. Soelemann, K. Roy and B. C. Paul, "Robust subthreshold logic for ultra-low power operation," *IEEE Transactions on VLSI Systems*, Vol.9, No.1 February 2001.

[2] T. S. Lande, D. T. Wisland, T. Saether and Y. Berg, "FLOGIC - Floating-Gate Logic For Low-Power Operation," ICECS '96, 1996.

[3] S. Aunet, Y. Berg, O. Tjore, Ø. Næss, T. Sæther "Four-MOSFET Floating-Gate UV-Programmable Elements for Multifunction Binary Logic," *Proceedings of the 5th World Multiconference on Systemics, Cybernetics and Informatics*, Vol.3, pp 141-144, Orlando, FL, USA, 2001.

[4] P. Hasler, T. S. Lande, "Overview of floating-gate devices, circuits and systems," *IEEE Transactions on Circuits and Systems - II: Analog and Digital Signal Processing*, Vol.48, No.1, January 2001.

[5] International Technology Roadmap for Semiconductors, Web documents http://public.itrs.net.

[6] T. Shibata and T. Ohmi, "A Functional MOS Transistor Featuring Gate-Level Weighted Sum and Threshold Operations," *IEEE Transactions on Electron Devices 39*, 1992.

[7] R. M. Swanson and J. D. Meindl, "Ion-implanted complementary MOS transistors in low-voltage circuits," *IEEE Journal of Solid-State Circuits, vol.7, no.2, pp.146-153, April 1972.*

[8] G. Schrom and S. Selberherr, "Ultra-low-power CMOS Technologies", *Int'l Semiconductor Conference, CAS*, Digest of Technical Papers, pp.237-246, 1996.

[9] H. Soelemann and K. Roy, "Ultra-low power digital subthreshold logic circuits," *ISPLED'99, Proc. of 1999 International symposium on Low Power Electronics and Design*, pp.94-96, San Diego, CA, USA, 1999.

[10] S. Aunet, "Real-time reconfigurable devices implemented in UV-light programmable floating-gate CMOS," *Ph.D. Dissertation 2002:52, Norwegian University of Science and Technology*, Norway, 2002.

[11] K. Rahimi, C. Diorio, C. Hernandez and M.D. Brockhausen, "A simulation model for floating-gate MOS synapse transistors," *Proc. of the IEEE Int'l Symposium on Circuits and Systems*, Vol.2, pp. 532-535, 2002.

[12] J. Alfredsson, S. Aunet and B. Oelmann, "Small Fan-In Floating-Gate Circuits with Applications to an Improved Adder Structure," Proceedinngs of IEEE VLSI 2007 Conference, Bangalore, India, 2007.

[13] J. Alfredsson, S. Aunet, "*Performance of CMOS and Floating-Gate Full-Adders Circuits at Subthreshold Power Supply"*, Proceedings of the International PATMOS Conference 2007, Sweden, September 2007.

[14] J. Rabaey, M. Pedram, P. Landman, "Low Power Design Methodologies" in J. Rabaey, M. Pedram (editors), *Low Power Design Methodologies*, Kluwer Academic Publishers, 1997.

[15] K. Granhaug and S. Aunet, "Improving Yield and Defect Tolerance in Multifunction Subthreshold CMOS Gates," Proceedings of the 21st IEEE International Symposium on Defect and Fault-Tolerance in VLSI Systems, DFT'06, pp.20-28, October 2006.

[16] P. Hasler, B. A. Minch and C. Diorio, "Floating-gate devices: they are not just for digital memories any more", *Proceedings of 1999 IEEE Int'l Symposium on Circuits and Systems*, Orlando, Florida, USA, 1999.

[17] J. Alfredsson, S. Aunet and B. Oelmann, "Basic speed and power properties of digital floating-gate circuits operating in subthreshold," *IFIP VLSI-SOC 2005, Proc. of IFIP International Conference on Very Large Scale Integration*, Australia, Oct 2005.

[18] A. Thomsen, M. A. Brooke, "A Floating-Gate MOSFET with Tunneling Injector Fabricated Using a Standard Double-Polysilicon CMOS Process," *IEEE Electron Device Letter*, Vol. 12, No.3,pp 111-113, 1991.

[19] D. D. Wen, "Design and Operation of a Floating Gate Amplifier," *IEEE Journal of Solid-State Circuits*, Vol. SC-9, No.6, December 1974.

[20] S.Aunet, Y. Berg, T. Ytterdal, Ø. Næss, T. Sæther, "A method for simulation of floating-gate UV-programmable circuits with application to three new 2-MOSFET digital circuits," *The 8th IEEE Int'l conference on Electronics, Circuits and Systems*, Vol.2, pp. 1035-1038, 2001.

[21] K. Granhaug and S. Aunet, "Six Subthreshold Full Adder Cells characterized in 90 nm CMOS technology," *Design and Diagnostics of Electronic Circuits and Systems*, pp.25 30, IEEE, April 2006

[22] Circuits Multi-Projects, CMP [online webpage] , http://cmp.imag.fr

[23] Y. Berg and T. S. Lande, "Programming Floating-Gate Circuits with UV-activated Conductances," *IEEE Trans. on Circuits and Systems - II: Analog and Digital Signal Processing*, Vol.48, No.1, pp. 12-19, 2001.

[24] E. J. Nowak, "Maintaining the benefits of CMOS scaling when scaling bogs down," IBM J. Res & Dev, vol.46, no.2/3, pp.169-180, 2002.

[25] K. Granhaug and S. Aunet, "Improving Yield and Defect Tolerance in Subthreshold CMOS Through Output-Wired Redundancy", Journal of electronic testing 2008.

[26] S. Aunet and H. K. Otnes Berges, "Statistical Simulations for Exploring Defect Tolerance and Power Consumption in 4 Subthreshold 1-bit Addition Circuits," Lecture Notes in Computer Science (4507) Springer 2007, ISBN 978-3-540-73006-4.

[27] S. Hanson et al, "Performance and Variability Optimization Strategies in a Sub-200mV, 3.5pJ/inst, 11nW Subthreshold Processor," IEEE Symposium on VLSI Circuits 2007, pp.152 - 153, June 2007.