Scaling Limits of Capacitorless Double Gate DRAM Cell

Nauman Butt and Muhammad A. Alam
School of Electrical and Computer Engineering
Purdue University
West Lafayette, IN 47906, U.S.A.
nbutt@purdue.edu

***Abstract*— In this paper, we consider the scaling of Capacitor-less Single Transistor (1T-0C) DRAM by classical (CL) and quantized-ballistic (QB) methods to establish that (1) it may be difficult to scale 1T-0C cell below 30nm channel length even with ultrathin (< 3nm) body because of the quantum confinement effects, (2) cumulative drain disturb time must be limited to ensure reasonable retention times, (3) the surround gate structures such as silicon nanowires (as 1T-0C cells) are expected to have more significant confinement effects, and (4) practical considerations such as the process variations in cell geometry and single events upsets are likely to remain important scaling concerns.**

Keywords-capacitorless; double gate; scaling; quantum confinement, simulation

I. INTRODUCTION

In classical DRAM cells, the logic level is indicated by charge (Q*)* stored on a capacitor node (1C) that is accessed by a pass transistor (1T). Since, Q = CV, high performance 1T-1C DRAM cells require larger capacitance (C) and higher voltage (V), opposite to the scaling trend of logic transistors. To remain relevant, DRAM designers have pioneered use of high-k dielectrics (e.g., Ta_2O_3, Al_2O_3, etc.), extreme capacitor geometry (e.g., trench, fin, stack, etc.), and innovative circuit techniques (e.g., hyperpage mode, etc.) [1]. Despite this comprehensive approach, however, the viability and scaling of DRAM capacitor below 65 nm node is not assured. A novel 1T-0C DRAM cell utilizes the floating body charging of an SOI transistor to replace the traditional DRAM capacitor and therefore is sometimes perceived to be more scalable than 1T-1C cells [2, 3]. For this DRAM cell, a logic state is defined by creating an excess or a shortage of the majority carriers inside the body of the transistor which cause a threshold voltage shift (ΔV_t) below or above its equilibrium value. The state is read by sensing the drain current which has two different levels corresponding to the two states. A high density 1T-0C DRAM would be very attractive for embedded DRAM applications as well as for the stand alone DRAM products.

Several designs of 1T-0C cells have been proposed which are based on both partially depleted (PD) and fully depleted (FD) SOI technologies [2-4]. Although removal of external capacitor reduces area overhead and reduces process complexity, the scaling of 1T-0C has its own specific challenges. For example, when the channel length (L_{ch}) is scaled down, the short channel effects like drain induced barrier lowering (DIBL) results in an increased junction leakage current which rapidly reduces the stored charge from the floating body. For a scalable design, the cell must have a strong gate control for which a surround gate or double gate (DG) geometry with ultathin body (t_{body}) is desirable. Smaller t_{body} not only reduces DIBL but also provides higher body coefficient – defined as the differential change of threshold voltage for a differential change in floating body potential. Despite these advantages, however, our analysis in this paper demonstrates that the quantum mechanical charge confinement in the body become significant at t_{body} < 3nm and offset the above mentioned advantages of small t_{body}. Thus the 1T-0C cell scaling is actually more difficult than perceived by the classical analysis, as discussed below. This paper is divided into five sections: Section II describes the cell operation and Section III explains the simulation models. Section IV illustrates the considerations for cell scaling and design. Finally, our conclusions are summarized in Section V.

II. CELL OPERATION

The **WRITE**, **READ** and **HOLD** operations of a floating body double gate 1T-0C cell are illustrated in Fig. 1. The two gates are connected to the world lines (WL1 and WL2) while the source/drain are connected to the bit lines (BL1 and BL2). The bias conditions for the lines during each of the cell operation are also shown. The bias conditions and doping are similar to those proposed in [2]. The state '1' is written by applying a high drain voltage (V_d) of 1.2V. The impact ionization creates electron-hole pairs near the drain. The electrons are rapidly swept out of the cell by the electric field while a portion of holes remains trapped in the body, increasing its concentration above the equilibrium value. The state '0' is written by applying a positive V_g and smaller V_d. The holes in the body are pushed out through

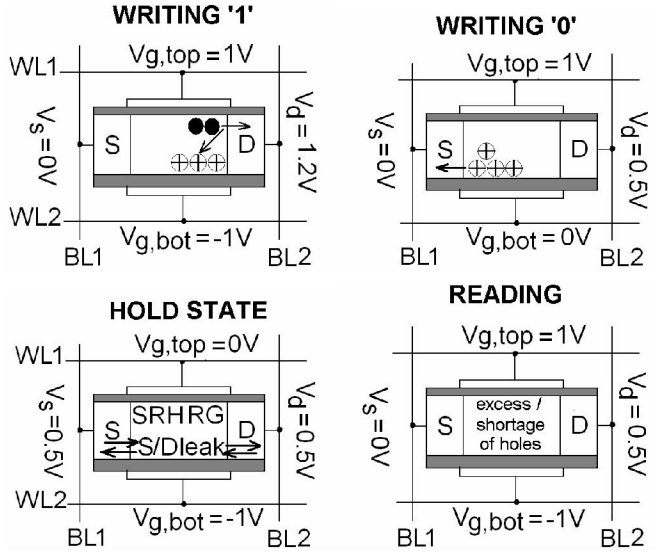

Fig. 1. The phenomena of **WRITE**, **READ**, and **HOLD** are illustrated with corresponding operating voltages. To write, $V_{g,top}$ is made high; state '1' is written by generating excess holes through impact ionization by applying V_d = 1.2V, state '0' is written by pulling holes out of body by applying small V_d. In hold state, barriers are created by applying negative $V_{g,bot}$; retention time is determined by SRH recombination generation rates and source/drain leakage. The threshold shift corresponding to excess/shortage of holes distinguishes the state during **READ**.

source /body junction and result in a shortage of holes in the body. This excess or shortage of holes modulates the floating body potential and consequently the threshold voltage (V_t). In the **HOLD** state, the gate voltages are reduced and small voltage is applied to the source and drain. This creates a large barriers at both the source and drain junctions of the body. The state written previously is retained for a refresh cycle. The retention time of the state is determined by SRH recombination/generation (RG) rate and the S/D leakage current. Finally, the state is read by sensing the drain current by applying a small V_d. The drain current has distinct levels for each of the two states because of the shift in threshold voltage.

III. SIMULATION MODELS

We use device simulator MEDICI to consider classical charge dynamics (CL model) and NANOMOS [5] to consider charge dynamics including the quantum quantization effects (QB model). In CL model, charge density is obtained by considering classical bulk conduction and valence bands and the drift-diffusion equation is solved for the carrier transport. The non-local model of impact ionization is included with soft threshold [6] and SRH R-G is considered with electron/hole lifetime of 1 microsecond. The band-to-band and trap-assisted tunneling are included in CL model but their contributions were found to be negligible for the doping and bias conditions used for the cell operations.

The QB model is based on NANOMOS which is a 2-D simulator for fully depleted thin-body DG MOSFET. It solves the Schrodinger equation in the transverse (thickness) direction self-consistently with the Poisson equation to get the quantized conduction sub-bands and the corresponding electron densities. The classical ballistic model of NANOMOS is used in which the ballistic transport of electrons in the channel direction is solved without including the quantum tunneling under the barrier. Our QB model also includes the quantized valence sub-bands (not implemented in NANOMOS) with the corresponding holes densities and the holes transport is treated classically. The charge dynamics are then obtained in the transient simulation which treat SRH and impact ionization similar to the CL model.

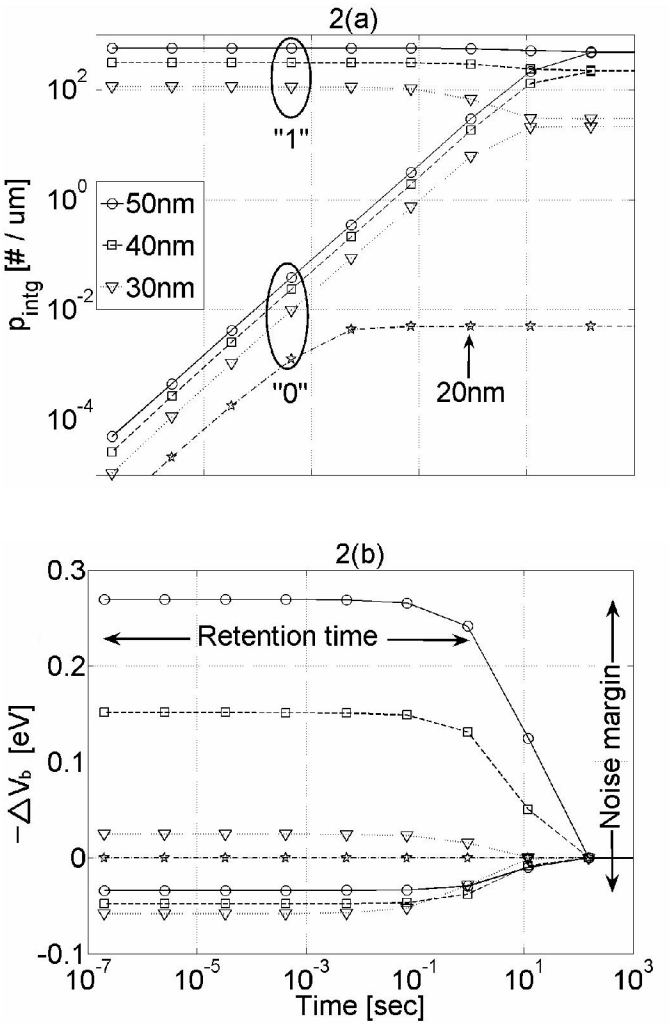

Fig. 2. P_{intg} (2a) and ΔV_b (2b) in **HOLD** state after **WRITE** '0' plotted as a function of **HOLD** time in 1T-0C cell of t_{body} = 10nm with various L_{ch}. The effect of DIBL is prominent in smaller L_{ch} where the ΔV_b gap becomes smaller and results in decreased sense margins.

IV. SCALING AND OPTIMIZATION OF 1T-0C CELL

A. Channel Length Scaling

Fig. 2 illustrates the retention time and the noise margin after writing a cell of t_{body} = 10nm for various L_{ch}. The retention time of a state is defined as the time for which the state can be correctly sensed by the **READ** operation after being written. The noise margin is the difference between the two states which can be defined in terms of difference in the threshold voltage (V_t) or the difference in the floating body potential (V_b). Fig. 2a shows the integrated holes density (P_{intg}), which is the total number of holes per unit width in the body, as a function of hold time after writing. As explained in section II, P_{intg} decreases during writing '0' and increases during writing '1'. In the **HOLD** state, SRH RG and the S/D junction leakage bring P_{intg} back towards its equilibrium value. The rates of these two phenomena determine the retention time. Fig. 2b shows the shift in the floating body potential (ΔV_b) during the hold time for the states. Since the impact ionization is not very high with V_d = 1.2V, ΔV_b caused by writing '0' primarily defines the noise margin. It should be noted that the noise margin becomes smaller for cells with smaller L_{ch}. In particular, ΔV_b goes to zero for L_{ch} = 20nm. This is the consequence of a well known short channel effect i.e., DIBL. In DIBL, the body potential comes under an increased influence of the source/drain field and the intrinsic holes concentration in the body is reduced. This can be seen in Fig. 2a where P_{intg} is decreasing with smaller L_{ch}. For $L_{ch} \leq 30$, P_{intg} is reduced to a level that is not enough to create a significant ΔV_b after a **WRITE** operation.

B. Body Thickness Scaling

There are two ways of reducing the effect of DIBL: (i) increased body doping concentration and (ii) reduced body thickness, t_{body}. But the increased body doping causes other two problems i.e, it increases the trap-assisted tunneling to the S/D and it increases the V_t-fluctuation due to random doping effects in ultrathin cells [2]. The former reduces the retention time and the latter shortens the noise margin. Reducing body-thickness, on the other hand, provides two advantages, i.e., it reduces DIBL an it increases the body coefficient ($\Delta V_t/\Delta V_b = -t_{ox,f}/t_{body}$, where $t_{ox,f}$ is the front oxide thickness) and hence a given ΔV_b can be translated into a bigger ΔV_t. However, cells with very small t_{body} suffer a radical drop in ΔV_b because of the quantum mechanical confinement of carriers. The splitting of the bands into quantized levels results in widening of the band gap and a reduction in the intrinsic carrier concentration. As a result, there are not enough holes in the body to cause a significant shift in V_b. Fig. 3 illustrates this in a cell with L_{ch} = 50nm and two different t_{body}. The conduction and valence bands (first conduction and valence subbands in QB model) along the channel direction are plotted along with the corresponding carrier densities for the two cases. For t_{body} = 5nm, both CL and QB models give almost the same results,

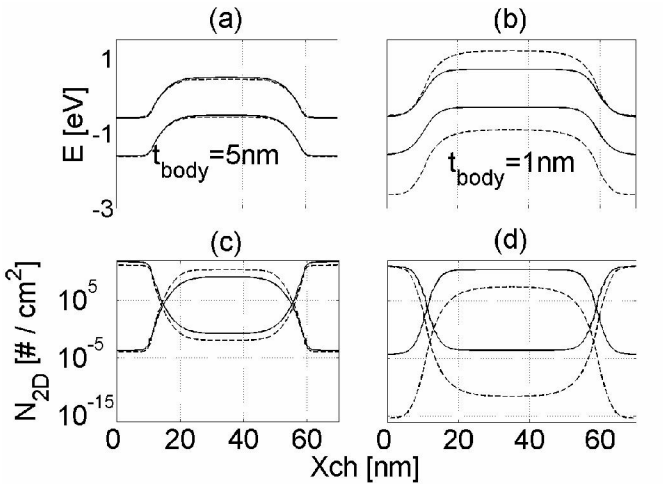

Fig. 3. The conduction & valence bands in the middle (t_{body} / 2) of cell with Lch = 50nm plotted along the channel direction for tbody = 5nm (left) & tbody = 1nm (right). Dotted and solid lines are result of QB & CL models respectively. The widening of the band gap can be observed with smaller tbody which decreases the equilibrium carrier concentration.

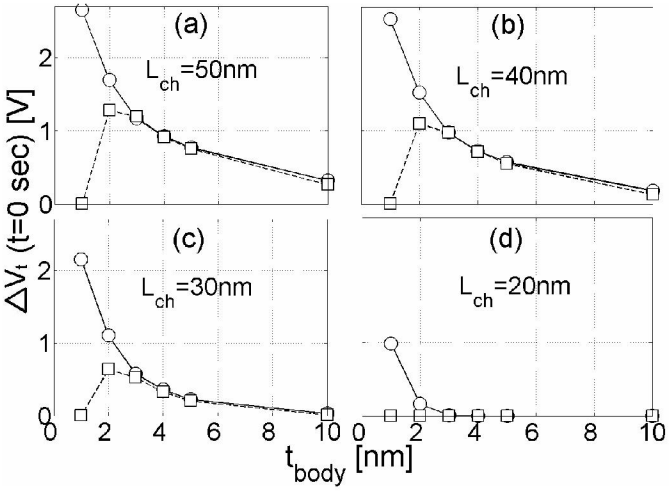

Fig. 4. Change in threshold voltage immediately after **WRITE** '0' for the cells with various dimensions obtained by QB (squares) and CL (circles) models. The models differ significantly at t_{body} < 3nm because of quantum mechanical confinement.

whereas in t_{body} = 1nm, the QB model shows a widening of the band gap and significant reduction in the carrier concentrations as compared to CL model. Fig. 4 further explains the effect of scaling t_{body} on the sense margin. The plots of ΔV_t at the time immediately after writing '0' are shown as a function of t_{body} for the cells of different L_{ch}. Both models give similar ΔV_t down to t_{body} = 3nm. In this range, both curves follow the expression: $\Delta V_t \sim t_{body}^{-1}$. For t_{body} < 3nm, the CL and QB models differ significantly because CL model continues to follow the above expression but QM model results in a much lower ΔV_t because of the quantum confinement effect. It can be noticed that QB

model predicts almost zero ΔV_t for L_{ch} = 20nm and t_{body} = 1nm.

C. *Design Space for the Scaled Cell*

Fig. 5 illustrates the design space for the dimensions of 1T-0C cell. The areas bounded by the allowed L_{ch} and t_{body} dimensions are shown as predicted by the two models while keeping $\Delta V_t \geq 0.3V$. The space above t_{body} = 3nm is limited by DIBL for L_{ch} < 40nm. Both models are consistent in this regime. For example, L_{ch} = 30nm has an upper limit of t_{body} = 4nm. For $t_{body} \leq$ 2nm, the design space from QB model is limited by the quantum confinement whereas the design space from CL model extends to 1nm. Thus, while CL model predicts that L_{ch} < 27nm could be achieved with t_{body} = 1nm, the QB model predicts that the scaling stops for L_{ch} < 27nm because the density of confined holes is so low that the cell volume does not have sufficient holes to support any **WRITE** operation.

D. *Other Considerations*

In practice, the drain disturb, process variations and radiation induced soft errors must also be considered in scaling of a 1T-0C cell. We explore the susceptibility of 1T-0C cell to drain disturb problem in Fig. 6 and find that the cumulative disturb time must be less than a millisecond for a given cell to hold its charge. This time sets the upper limit of number of write operations in the neighboring cells which share a common drain bit-line before the cell in **HOLD '0'** must be refreshed.

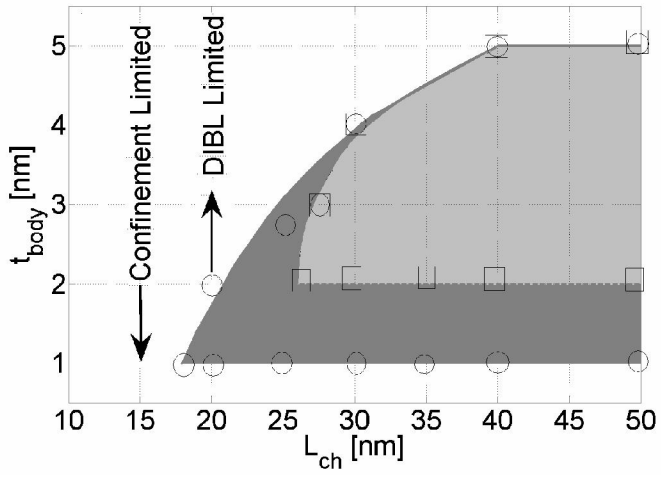

Fig. 5. The allowed design space of 1T-0C cell when dimensions are scaled while keeping $\Delta V_t \geq 0.3V$. The whole area bound by outer curves (circles) is the result of CL model and the lightly shaded area is from QB model (squares). The dark shaded area is where the two models differ due to the effect of charge confinement. At higher t_{body}, DIBL limits the design space while at $t_{body} \leq$ 2nm, quantum confinement limits the design.

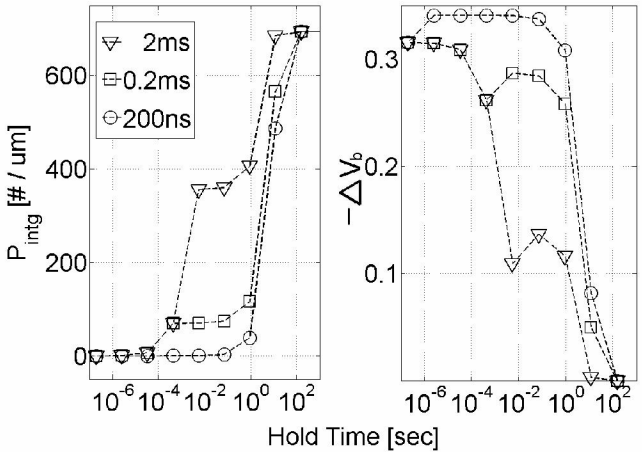

Fig. 6. The effect of **WRITE** '1' drain disturb on 1T-0C cell in **HOLD** '0' with L_{ch} = 50nm and t_{body} = 10nm. P_{intg} (left) and - ΔV_b (right) are plotted as a function of hold time for cumulative disturb times of 200ns, 0.2ms and 2ms with higher values reducing sense margin and retention time.

V. CONCLUSIONS

In this paper, we have systematically analyzed the scaling limits of a novel; (presumably) more scalable double gate 1T-0C cell to find that in practice the scaling of the cells may actually be more difficult than previously presumed. In contrast to the analysis based on the classical model, the model including the quantum confinement of carriers predicts that the confinement effect causes a significant loss of sense margin between the two states of 1T-0C cell for ultra-thin body (< 3nm). For a slightly greater body thickness, both models predict that DIBL limits the channel length scaling. The trade-off between confinement, body coefficient and DIBL determines the limits of 1T-0C cell scaling. In practice, other considerations like drain disturb, process variations and single event upset may further restrict scaling. The surround gate geometry for 1T-0C cell is expected to be less scalable because of two dimensional confinements. All these have significant implications for 1T-0C cell as a potential replacement for classical 1T-1C DRAM.

REFERENCES

[1] J. A. Mandelman, R. H. Dennar, G. B. Bronner, J. K. DeBrosse, R. Divakaruni, Y. Li & C. J. Radens, *IBM J. Res. & Dev.*, Vol. 46, No. 2/3, 2002

[2] C. Kuo, T. King and C. Hu, *IEEE Transactions on Electron Devices*, Vol. 50, No. 12, Dec. 2003

[3] E. Yoshida, T. Miyashita, T. Tanaka, *IEEE Electon Device Letters*, Vol. 26, No. 9, Sep. 2005

[4] S. Okhonin, M. Nagoga, J. M. Sallese, and P. Fazan, *IEEE Intl. SOI Conf.*, 2001, pp. 153-154

[5] Z. Ren, R. Venugopal, S. Goasguen, S. Datta, M. Lundstrom, "nanoMOS2.5: A Two -Dimensional Simulator for Quantum Transport in Double-Gate MOSFETs,"IEEE Trans. Electron. Dev., and IEEE Trans. on Nanotechnology, joint special issue on Nanoelectronics, Vol. 50, pp. 1914-1925, 2003.

[6] C.Jungemann, R.Thoma and W.L.Engl, *Solid-StateElectronics*, Vol. 39, No. 7, pp.1079-1086, 1996.

