Extended analysis of the $Z^2$-FET: Operation as capacitor-less eDRAM

Carlos Navarro, Joris Lacord, Mukta Singh Parihar, Fikru Adamu-Lema, Meng Duan, Noel Rodríguez, Binjie Cheng, Hassan El Dirani, Jean-Charles Barbe, Pascal Fonteneau, Maryline Bawedin, Campbell Millar, Philippe Galy, Cyrille Le Royer, Siegfried Karg, Paul Wells, Yong-Tae Kim, Asen Asenov, Sorin Cristoloveanu and Francisco Gamiz

Abstract—The $Z^2$-FET operation as capacitor-less DRAM is analyzed using advanced 2D TCAD simulations for IoT applications. The simulated architecture is built based on actual 28 nm FD-SOI devices. It is found that the triggering mechanism is dominated by the front-gate bias and the carrier’s diffusion length. As in other FB-DRAMs, the memory window is defined by the ON voltage shift with the stored body charge. However, the $Z^2$-FET’s memory state is not exclusively defined by the inner charge but also by the reading conditions.

Keywords—1T-DRAM, capacitor-less, feedback effect, fully depleted, ground plane, lifetime, sharp switch, SOI, $Z^2$-FET.

I. INTRODUCTION

Among other uses such as ESD [1], the $Z^2$-FET or similar structures, can be considered as a promising candidate to operate as single-transistor DRAM cell [2], [3]. In order to succeed, 1T-DRAM cells need to comply with several requirements [4]: i) low operating voltage and power consumption, ii) long retention time, iii) high density integration, iv) fast random access, and v) easy fabrication.

The $Z^2$-FET features: i) possibility of employing ultra-thin SOI films without suffering from the supercouppling effect [5], ii) high retention times [6], iii) large current ratio, iv) regenerative reading [2] and v) CMOS flow compatibility. These advantages motivate an in-depth study of the cell to fully understand the operation mechanisms occurring inside the device enabling an efficient optimization, hence boosting the performance.

II. SIMULATION SETUP: STRUCTURE AND MODELS

To explore the $Z^2$-FET operation, 2D numerical simulations were carried out by employing Synopsys TCAD tool [7]. The default structure in Fig. 1 corresponds to experimental $Z^2$-FET devices in 28 FDSOI technology [8]. The $Z^2$-FET consists of a FD (Fully Depleted) SOI gated P-I-N diode where the front gate does not cover the entire Silicon body spanning from the cathode to the anode. This asymmetric architecture leads to two different regions characterized by their corresponding lengths: the portion where the front-gate is present, noted $L_G$, and the ungated region $L_{1n}$, Fig. 1.

Room temperature (300 K), Fermi-Dirac statistics, density gradient quantization [7], Shockley-Read-Hall (SRH), surface recombination and band-to-band tunneling [9] models were included. The carrier mobility accounts for the doping dependence, high field velocity saturation, transverse electric field (with remote-Coulomb scattering) and Thin-Layer mobility model [7]. No impact ionization models were included since the $Z^2$-FET is insensitive to them [2].

In order to obtain similar DC hysteresis as the experimental data, two parameters were modified from default values:

1) The anode and cathode access resistance enabled the fitting of the curves at high lateral fields ($V_A >> 0$ V). For simplicity, identical values were considered at both terminals, $R_{A/K} = 400 \, \Omega \cdot \mu m$.

2) The maximum carrier lifetime for electrons and holes.

The final values were $\tau_n = 2.5 \cdot 10^{-8}$ s and $\tau_p = 10^{-8}$ s.

Figure 2 illustrates the current-voltage characteristics from a) experimental and b) simulation data. The hysteresis is well reproduced together with the maximum current value.
Fig. 2. a) Experimental and b) 2D simulated $I_A(V_A)$ hysteresis characteristics for several front-gate voltages. The applied anode bias pattern consists on a triangular signal from 0 to 1.5 V with 24.16 s ramping (rising/falling) time. Default parameters with $L_G = L_{In} = 200$ nm, $V_K = 0$ V and $V_{BG} = -1$ V.

Fig. 3. a) Energy bands edge diagram for $V_A = 0$ V and b) anode current curves when biasing the gates ($Z^2$-FET operation in solid lines, $V_{FG} = +1$ V and $V_{BG} = -1$ V) or not (diode operation in dashed lines, $V_{FG} = 0$ V and $V_{BG} = 0$ V).

III. Z$^2$-FET BASICS

In order to forward bias the N-type $Z^2$-FET, the $N^+$ cathode (K) is grounded while a positive voltage is applied to the $P^+$ anode (A). By default, at $V_{FG} \approx V_{BG} \approx 0$ V, the device behaves as a typical P-I-N diode. The $Z^2$-FET basic idea is to induce a complementary energy barrier along the intrinsic body transforming the intrinsic region of the diode into a P-N junction as a result of the front (positive) and back-gate (negative) biasing. This bias finally leads to a N-P-N-P virtually-doped structure. Figure 3a shows the energy bands edge diagram when inducing (solid lines) or not (dashed lines) the virtual doping in the $Z^2$-FET body. Figure 3b illustrates the stationary $I_A(V_A)$ curves for the two previous gate biasing scenarios. A triangular anode signal from 0 to 1.5 V is applied with a ramping (rising/falling) time of several tens of seconds. With gate biasing the energy barriers are large and the device presents hysteresis with a sharp current switch (solid line). Otherwise the diode behavior is observed (dashed line). The abrupt changes in the conductance can be explained as follows.

A. Shockley diode analogy

The $Z^2$-FET resembles a Shockley diode [10] since it features three homo-junctions (J1-3 in Fig. 3a) without a gate current terminal as thyristors [11]. The main difference is that the junctions are formed by physical doping (implants) instead of gate-induced “virtual doping” which completely modifies the operation and switching.

The N-P-N-P structure can be also seen as two bipolar transistors (one N-P-N and another P-N-P) driving each other into saturation [11]. This mechanism is regenerative and causes the increase in the carrier injection from the lateral terminals. When the positive feedback loop grows enough to forward bias J2 (both BJT turn into saturation abruptly), the positive feedback effect becomes unstable triggering the device ON.

B. Detailed explanation of barrier modulation

Figure 4 shows the evolution of the $Z^2$-FET carrier profiles, energy bands and quasi-fermi energy levels (QFE, $E_{F_{n,p}}$) with time ($\approx 0.1$ ms step). A very long (seconds) pulse of 0.85 V with 10 µs rising time is applied to the anode ($V_{FG} = 1$ V, $V_{BG} = -1$ V and $V_K = 0$ V). The bias values were selected to illustrate the device turn-on rather than to optimize the memory retention. The device current readout is depicted in the inset of Fig. 4a. It exhibits the sharp switch at around 2.8 ms. This abrupt switch is induced by a positive feedback effect between the carrier injection from anode and cathode. However, the barriers collapse is induced by the quasi-Fermi energy shift. It is worth noting that this effect is not a transient effect since it is also observed in DC stationary state. When the anode pulse is applied, the concentration of holes rises with time due to the reduction in the J1 built-in barrier, Fig. 4a. However, before the triggering, the hole current density increase is slower, Fig. 4c. If the hole concentration ($p$) rises faster than the current density ($J_p$), the hole QFE gradient needs to be reduced according to Eq. 1.

$$J_p = \mu_p \cdot p \cdot \nabla E_{F_p}$$

where the hole mobility $\mu_p$ is almost constant. As the hole QFE is pinned by the anode energy, the decrease of the hole QFE gradient implies a downward shift in the ungated region, Fig. 4d. Given that the hole population does not increase abruptly in the ungated region (Fig. 4a), the distance of the hole QFE with the valence band edge remains almost constant with time, Fig. 4f. If the hole QFE moves downward and the distance remains constant, the valence band, thus also the conduction band, must follow this shift, Fig. 4d. This explains the cathode-body (J3) barrier decrease. The same mechanism happens for electrons when the J3 barrier decreases. They reach the gated region further reducing the anode-body (J1) barrier due to the electron QFE increase (as in Eq. 1), Fig. 4b-e. Once the feedback mechanism goes unstable (i.e., the injection of carriers is high enough), the $Z^2$-FET triggers.

The $Z^2$-FET sharp switch can happen in both the forward and backward directions or exclusively in the forward sense. The absence of the abrupt current drop implies that J3 is still in forward biased mode even when $V_A$ is strongly reduced.
Fig. 4. a) Hole and b) electron horizontal carrier profiles as a function of time. The inset represents the current readout with the sharp switch at $t \approx 2.8$ ms. c) Absolute horizontal (x-sense) hole current density. Energy band edges with d) hole and e) electron QFE. f) Distance from valence band edge to hole QFE (extracted from d at $x = -0.225 \mu m$). The time step between curves corresponds to $\approx 0.1$ ms. Profiles extracted 0.1 nm away from the top-gate oxide interface. $V_A = 0.85$ V, $V_{FG} = 1$ V, $V_{BG} = -1$ V and $V_K = 0$ V.

**IV. TIME-DEPENDENT OPERATION**

In this section the transient behavior of the $Z^2$-FET is analysed when operating as a memory.

**A. Dynamic current states**

The $Z^2$-FET memory operation is based on a transient shift of the ON triggering voltage thanks to the population (high current ‘1’-state) or depletion (low current ‘0’-state) of the body. By default, the body is populated making the ‘1’-state stable. Figure 5 depicts the forward $I_A(V_A)$ characteristics after a 10 ns long programming for the a) ‘0’ and b) ‘1’ states. The reading access speed is approximated by using the ramping time. The $V_{ON}$ after $W_0$ is larger than after $W_1$. This transient shift enables the memory window, i.e., useful anode voltages to sense the memory state, defined as the bias range in-between both curves. An example of memory window is depicted in Figure 5c for a 1 ms ramping time. The $V_{ON}$ approaches until they finally merge preventing the memory operation. Figure 6a shows the $Z^2$-FET response to a constant reading for different anode voltages after depleting the body ($W_0$). Depending on the anode voltage, the device presents three different behaviors:

1) For $V_A \leq 0.845$ V the device is never switched ON no matter the elapsed time. The $J_1$ built in energy barrier prevents the injection of sufficient holes to start forward biasing $J_3$. Even if a certain amount of holes gets into the body, the leakage at $J_3$ is higher or equal so that there is no net charge increase and the hole density remains constant with time. Hence, the positive feedback mechanism
concentrations with time below the front-gate (mid-channel, voltage at different times extracted from a). c) Hole and d) electron density time after depleting the body. b) Anode current as a function of anode voltage at different times: i) stationary DC conditions (circles, \( t = 0 \text{ms} \)), ii) immediately after the rising pulse \( V_A = 0.85 \text{V} \) (no triggering) and \( V_A = V_{ON-DC} = 0.85 \text{V} \) (triggering). Values extracted 0.1 nm away from the top-gate oxide interface. \( V_{FG} = 1 \text{V} \), \( V_{BG} = -1 \text{V} \) and \( V_K = 0 \text{V} \).

never starts and the current remains low, '0'-state. Figure 6c-d shows the horizontal carrier profiles for holes and electrons against time at mid-channel \( (x = 0 \mu \text{m}) \) and in the intrinsic region \( (x = -0.225 \mu \text{m}) \) for \( V_A = 0.845 \text{V} \) (no triggering) and \( V_A = V_{ON-DC} = 0.85 \text{V} \) (triggering). Values extracted 0.1 nm away from the top-gate oxide interface. \( V_{FG} = 1 \text{V} \), \( V_{BG} = -1 \text{V} \) and \( V_K = 0 \text{V} \).

2) On the other hand, if \( V_A \geq 0.99 \text{V} \), the anode injection of holes is always large enough to trigger the device during the anode bias pulse rising time, i.e., high current '1'-state is always read, regardless of the reading pulse width.

3) The switching from the '0' to the '1'-state only appears after a given time for the anode bias range in-between \( (0.845 < V_A < 0.99 \text{V}) \). The injection is not very high but exceeds the J3 leakage current so the charge stored is gradually increased, hence triggering is finally achieved.

Figure 6b evidences the ON voltage shift with time. It illustrates the anode current as a function of the reading anode bias at different times: i) stationary DC conditions (circles, \( t \to \infty \)), ii) immediately after the rising pulse edge \( (t = 0.01 \text{ms}, \text{ squares}) \) and iii) at an intermediate time \( (t = 0.1 \text{ms}, \text{ triangles}) \). Before applying the anode reading pulse the device body is depleted, '0'-state, with high energy barriers. As time goes by, carriers are injected or generated reducing the barriers and the ON voltage toward the DC steady curve. Figure 6 also shows that, contrary to other floating-body memory cells [12], [13] where the read memory state is fully defined by the charge stored in the body, the state read in the \( \text{Z}^2\)-FET memory relies also upon the carrier injection from the lateral terminals. If the carrier injection while reading is not sufficient to trigger the diode current, the device presents low-conductivity ('0'-state) even if the device is fully populated with electrons. The carrier injection depends on two conditions:

1) The previous state of the device, i.e. the charge stored inside the device prior to the reading. The body charge, in particular, underneath the gated region, defines the initial energy barrier height between anode and channel before the injection takes place. Lower concentrations of carriers, with respect to the stationary state, enable higher energy barriers, as in the deep depletion mode of a MOS capacitor [14]. Higher barriers reduce the injection from the lateral terminals into the body, even if the reading conditions \( (V_A \text{ and pulse duration}) \) are identical. If no charge is stored under the gate, the reading anode pulse does not produce the triggering ('0'-state). On the other hand, if many electrons are stored, the application of the same biasing switches ON the device ('1'-state), Fig. 7. The charge modulation inside the body, i.e. memory state programming, is as follows:

a) To enhance the carrier population \( (W_1) \), the device is switched ON (P-I-N diode mode) increasing the anode voltage while grounding the front-gate bias. When the front-gate bias returns to a positive voltage, the induced potential well collects the available carriers.

b) The evacuation of charge \( (W_0) \), carrier depletion, is achieved by capacitive coupling: the front-gate is grounded to eliminate the potential well and evacuate all subsisting electrons.

2) The anode voltage and reading pulse width define the overall amount of carriers injected in the body enabling or not the sharp switch to occur. The reading conditions
also affect the diffusion of holes into the body of the transistor. Higher anode voltages reduce the anode-body junction built-in barrier enhancing the injection. Furthermore, longer reading pulses also enlarge the total amount of carriers getting into the body. This implies that similar barrier conditions, i.e. charge stored within the device, could lead to different memory states depending on the reading parameters. In the $Z^2$-FET, the anode terminal combines the roles of gate and drain in other capacitor-less cells. Figure 8 shows the a) bias pattern ($W_0$-$R$-$W_1$-$R$ sequence) and b) current readout for different anode voltages and reading times. Low anode voltages are not enough to adequately switch ON the device even after programming the '1'-state ($V_A \leq 0.8$ V). In contrast, high anode biases ($V_A \geq 1.2$ V) always trigger the device regardless of the memory state previously programmed. Narrow reading pulses solve this problem by reducing the injected charge as shown in Fig. 8c for $t_{\text{Reading}} \leq 50$ ns (not affecting the '1'-state, Fig. 8d).

### B. Hysteresis cycle

The hysteresis window shown in Fig. 2a does not describe the memory operation. Nevertheless, its analysis provides useful information on the sensitivity of other parameters:

1. **Ramping time:** Faster ramps/pulses reduce the amount of injected carriers, hence $V_{ON}$ increases, Fig. 9a. Long ramping times ($\geq 0.1$ s), yield the stationary DC biasing simulations in Fig. 3b (solid line).

2. **Carrier lifetime / device length:** The lifetime of minority carriers, $\tau_{n,p}$, together with the diffusion coefficient, $D_{n,p}$ (mobility, $\mu_{n,p}$, temperature, $T$, and the thermal voltage, $k \cdot T/q$), defines the diffusion length [15]:

$$L_{n,p} = \left( D_{n,p} \cdot \tau_{n,p} \right)^{1/2} = \left( \mu_{n,p} \cdot \frac{k \cdot T}{q} \cdot \tau_{n,p} \right)^{1/2} \quad (2)$$

$L_{n,p}$ indicate the average distance electrons and holes travel before they recombine. The electron diffusion length, $L_n$, should be compared to $L_{In}$ (where electrons are minority carriers) and $L_p$ compared to $L_G$ (where holes are minority carriers). This comparison sets the effective injection into the body from cathode and anode, respectively. If these regions are too long compared to $L_{n,p}$, most of injected carriers do not cross the gate/intrinsic regions and do not contribute to the barriers collapse. The effect is similar to that of the base width.
in a BJT transistor. Short carrier lifetimes reduce this diffusion length, hence the effective injection. As a result, $V_{ON}$ shifts to larger anode biases, Fig. 9b. Similarly, long devices exhibit larger $V_{ON}$, Fig. 9c. Note also that in long devices the barriers are stronger, free from short-channel effects.

3) Temperature: Low temperatures reduce the carrier energy [16], thus depress the amount of carriers able to surmount the J1/J3 barriers. Furthermore, the diffusion length is also reduced. As a result, the effective injection decays and $V_{ON}$ increases, Fig. 9d. The retention time and the base current level are also modified by temperature-controlled generation and parasitic carrier injection.

4) Front-gate bias / top-oxide thickness: $V_{FG}$ induces the potential well to gather electrons below the top-oxide. Larger biases induce higher energy barriers reducing the hole injection. Hence, larger anode voltages are required to trigger the device, Fig. 9e. The sensitivity to this parameter can be enhanced by thinning the top front-oxide.

5) Back-gate bias / buried-oxide thickness: The back-gate bias induces a well for holes but has a parasitic influence in the gated region weakening the front-gate barrier. More negative $V_{BG}$ reduce the $V_{ON}$, Fig. 9f.

C. Cell optimization guideline

The improvement of the memory cell’s figures of merit represents a challenge since many parameters are inter-related and affect the cell in different manners. For instance, large anode voltages boost the current margin, $I_1 - I_0$, but degrade the ratio, $I_1/I_0$, and retention time. The previous trade-off can be partially solved if controlling the anode injection by rising the top-gate bias so that $V_A < V_{FG}$ is satisfied. This enhances the hole barrier reducing $I_0$ while improving the retention. An increase of $V_{ON}$ and the power consumption are expected though. The injection can be also restricted by enlarging the $Z^2$-FET length (which also reinforces the barrier limiting SCE) or by reducing the diffusion length (rising $\tau_{n,p}$ via surface preparation or body doping, or decreasing the mobility by adapting the crystallographic orientation or using mobility-degrading strain). The smallest functional back-gate bias is suggested to sustain the front-gate impact and to enable simpler design and lower energy usage. Finally, use as narrow reading pulses and as fast ramping times as possible to minimize the ‘0’-state degradation.

Table I summarizes the impact of these parameters on the $Z^2$-FET behavior as eDRAM cell. It is worth noting that the same trend is found for the $V_{ON}$ and the retention time.

### Table I. Impact of Key Parameters on the $Z^2$-FET Operation As 1T-DRAM

<table>
<thead>
<tr>
<th>Parameter</th>
<th>$V_{ON}$</th>
<th>Current Ratio</th>
<th>Retention Time</th>
<th>Related to</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ramping Time</td>
<td>$\tau_{n,p}$</td>
<td>$I_{ON}/I_{OFF}$</td>
<td>$T$ and $\mu_{n,p}$</td>
<td>$T$, $I_{ON}$, $I_{OFF}$, $\mu_{n,p}$</td>
</tr>
<tr>
<td>$L_{SB}$, $L_{FG}$</td>
<td>$\tau_{n,p}$</td>
<td>$I_{ON}/I_{OFF}$</td>
<td>$T$ and $\mu_{n,p}$</td>
<td>$T$, $I_{ON}$, $I_{OFF}$, $\mu_{n,p}$</td>
</tr>
<tr>
<td>$V_{FG}$</td>
<td>$\tau_{n,p}$</td>
<td>$I_{ON}/I_{OFF}$</td>
<td>$EOT$, $V_{BG}$</td>
<td>$EOT$, $V_{BG}$</td>
</tr>
<tr>
<td>$V_{BG}$</td>
<td>$\tau_{n,p}$</td>
<td>$I_{ON}/I_{OFF}$</td>
<td>$EOT$, $V_{BG}$</td>
<td>$EOT$, $V_{BG}$</td>
</tr>
</tbody>
</table>

### Conclusion

The $Z^2$-FET DC and memory operations have been investigated by 2D TCAD simulations. The basics and the origin of the sharp switch are reexamined. $Z^2$-FETs exhibit a complex behavior where the memory state depends on the device architecture and on the reading conditions. The impact of several metrics is quantified and explained while tips to improve the performance as eDRAM are provided.

### References


