High-Density 4T SRAM Bitcell in 14-nm 3-D CoolCube Technology Exploiting Assist Techniques

Reda Boumchedda, Jean-Philippe Noel, Bastien Giraud, Kaya Can Akyel, Mélanie Brocard, David Turgis, Edith Beigné

High-Density 4T SRAM Bitcell in 14-nm 3-D CoolCube Technology Exploiting Assist Techniques

Réda Boumchedda, Jean-Philippe Noel, Bastien Giraud, Kaya Can Akyel, Mélanie Brocard, David Turgis, and Edith Beigne

***Abstract*— In this paper, we present a high-density four-transistor (4T) static random access memory (SRAM) bitcell design for 3-D CoolCube technology platform based on 14-nm fully depleted - silicon on insulator MOS transistors to show the compatibility between the 4T SRAM and the 3-D design and the considerable density gain that they can achieve when combined. The 4T SRAM bitcell has been characterized to investigate the critical operations in terms of stability (retention and read) taking into account the postlayout parasitic elements. Thus, failure mechanisms are exposed and explained. Based on this paper, a data-dependent dynamic back-biasing scheme improving the bitcell stability is developed. A specific read-assist circuit is also proposed in order to enable a large number of bitcells per column in a memory array. Finally, the designed bitcell offers up to 30% area gain compared to a planar six-transistor SRAM bitcell in the same technology node.**

***Index Terms*— 3-D monolithic, back bias, fully depleted - silicon on insulator (FD-SOI), read assist, static random access memory (SRAM).**

I. Introduction

OVER the past few decades, semiconductor industry has pushed for innovations on both circuit design and manufacturing technology in order to follow Moore's law, which states that the transistor density must double every new technology node. This later comes with the main challenge of increasing the performances, while reducing the power consumption [1], [2]. Nowadays, we face major issues since the size of transistors are reaching subatomic level, which leads to the rise of inevitable parasitic phenomena such as short-channel effects [3], process variability [4], and increased leakage [5], [6]. As a consequence, different stacking methods are developed in order to overcome scaling down bottlenecks in conventional planar design and to allow higher densities for system-on-chips (SoCs). This trend can be compared to an architect constructing a building by stacking flats to use efficiently a small land. In microelectronics, the equivalent is designing SoCs by stacking circuits or transistors to increase density for the same 2-D footprint. This is called 3-D design, and the manufacturing is generally named 3-D integration.

The 3-D integration has first appeared at die level by stacking vertically multiple dies [7]. The 3-D integration at the MOS level has appeared in the early 2000, and it is named the 3-D sequential integration or 3-D monolithic. In this technology, several layers of transistors with their metallization are processed sequentially and connected with via as small as the ones of the back end of line (BEOL). This would offer the highest density of devices and vertical connections among other 3-D technologies. This leads to a significant enhancement in performances and in power consumption thanks to the wire length reduction [8]. The main challenge of this technology consists in the complexity of the manufacturing process, since it requires a low-temperature process to make the upper devices in order to minimize the thermal budget of the chip, i.e., not destroying the metallization and devices in the bottom level [9]. Among the 3-D monolithic technologies, 3-D sequential LETI CoolCube technology [10] offers a very fine 3-D interconnect pitch compared to existing technologies and opens the way for efficient 3-D-VLSI circuits with the hope of reducing the congestion on the BEOL while providing real 3-D routing possibilities [11].

Static random access memory (SRAM) occupies the most significant area on SoC, and thus a lot of effort is put in the semiconductor industry to keep the SRAM in Moore's roadmap. A straightforward way of reducing the overall SRAM area is to reduce the number of transistors in the SRAM bitcell. Besides the well-known six-transistor (6T) SRAM bitcell, the four-transistor (4T) SRAM bitcell [12], [13] was first introduced in 1970 and it was commonly used in stand-alone SRAM devices. Compared to the 6T SRAM bitcell, the pull-up (PU) pMOS transistors were replaced by resistors with high resistances. It provided an advantage in density at the expense of manufacturing complexity since an extra layer of polysilicon was used to implement the PU resistors. The principal drawback of that design is the considerable increase in static power due to the constant current flow through the resistor and the pull-down nMOS. The interest for 4T SRAM bitcell has been lost as it appeared to be very difficult to scale below 1.8-V power supply due to its poor stability [14].

Manuscript received September 15, 2016; revised December 12, 2016 and February 13, 2017; accepted March 15, 2017. Date of publication April 19, 2017; date of current version July 24, 2017. *(Corresponding author: Réda Boumchedda.)*

R. Boumchedda is with STMicroelectronics, 38926 Crolles, France, with the University of Grenoble Alpes, F-38000 Grenoble, France, and also with the CEA-LETI, MINATEC Campus, F-38054 Grenoble, France (e-mail: boumchedda.reda@gmail.com).

J.-P. Noel, B. Giraud, K. C. Akyel, M. Brocard, and E. Beigne are with the University of Grenoble Alpes, F-38000 Grenoble, France, and also with the CEA-LETI, MINATEC Campus, F-38054 Grenoble, France

D. Turgis is with STMicroelectronics, 38926 Crolles, France.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TVLSI.2017.2688862

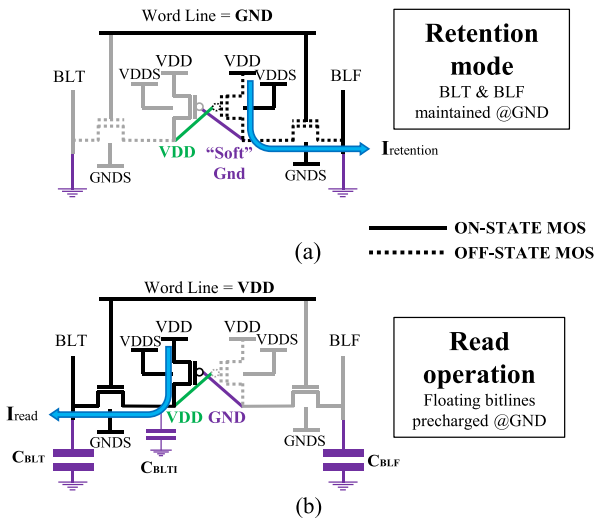

Fig. 1. Critical path of (a) retention and (b) read operations of 4T DL SRAM bitcell.

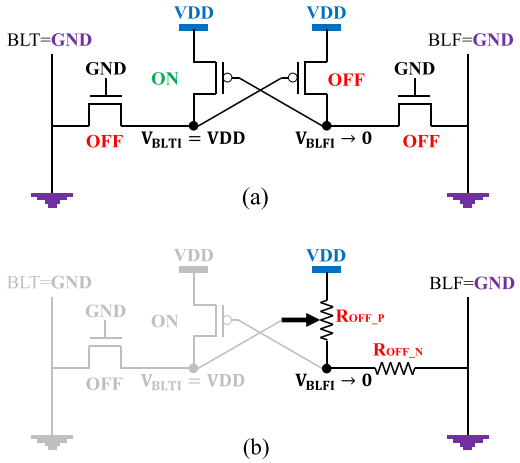

Fig. 2. (a) MOS based and (b) simplified models of the critical path during the retention of 4T DL SRAM bitcell.

Thanks to the fully depleted-silicon on insulator (FD-SOI) technology [15], the 4T SRAM bitcell can again offer good stability through threshold voltage (V_T) modulation using back biasing, and this can be realized without changing the regular process and without increasing the static power consumption. Moreover, since the 4T bitcell possesses two pMOS transistors and two nMOS transistors, thus an equal number of pMOS and nMOS, it is a promising candidate for the sequential 3-D technology (pMOS bottom layer and nMOS top layer) compared to the 6T bitcell, in which the asymmetry in the number of nMOS and pMOS limits the density gain [16]. Furthermore, the 3-D design could improve the performance and consumption of 4T SRAM and make it equal or even better than the planar 6T SRAM.

The reminder of the paper is organized as follows. Section II exposes 4T bitcell failure mechanisms. Section III details the 4T bitcell architecture in 3-D CoolCube technology. Section IV presents the in-house testbench and the simulation results. Section V demonstrates the application of a dynamic back bias to strengthen the stability of the 4T SRAM bitcell. Section VI proposes a powerful read-assist (RA) technique that allows a large number of bitcell per column (b/c). Finally, Section VII draws conclusions.

II. 4T SRAM Bitcell Failure Mechanism

In this section, the failure mechanisms in retention mode and read operation of the 4T SRAM bitcell are deeply analyzed to understand why those failures occur and justify the pertinence of the solutions that we provide in this paper. The 4T driver less (DL) architecture is chosen for this study, but this work can also be applied on the 4T load less (LL, no PU pMOS) [17]. This choice was motivated by the retention robustness of the 4T DL compared to the 4T LL based on the current data of 14-nm 3-D CoolCube technology.

While the 6T bitcell has its internal nodes firmly maintained at the supply voltages (V_{DD} and G_{ND}), the 4T bitcell has one of its node maintained through an "equilibrium" of leakage current and this yields to a potential retention problem. The critical path for the retention of the 4T DL bitcell is shown in Fig. 1(a) as the current flows from V_{DD} to BLF passing through the internal node set to "soft" G_{ND} (i.e., not physically maintained at G_{ND} by an ON-state nMOS). If the current leakage of the nMOS [pass gate (PG)] is lower than the current leakage of the pMOS (PU), the internal node voltage will increase, and if the internal node voltage gets too close to V_{DD}, the bitcell will toggle and the stored data will be lost. To have a robust retention, the nMOS PG must be more "leaky" than the pMOS PU, which results in a difference of resistances; this later is linked to the threshold voltage (V_T) in the cutoff region of the two MOS.

The critical path of the retention can be modeled as a voltage divider by replacing the nMOS and pMOS by their equivalent resistance (respectively, R_{OFF_N} and R_{OFF_P}) as shown in Fig. 2. The value of the internal node can be described as shown in the following:

$$V_{BLFI} = \frac{R_{OFF\_N}}{R_{OFF\_N} + R_{OFF\_P}} \times V_{DD}. \quad (1)$$

If we want V_{BLFI} to be as close as possible to G_{ND}, the ratio of resistances given by (1) must be as close as possible to "0." This can be read as

$$\boxed{R_{OFF\_P} \gg R_{OFF\_N}}. \quad (2)$$

Fig. 3 presents the ratio of resistances that defines the internal node voltage. We note that the retention is more stable at high temperatures. Monte Carlo (MC) simulation with 1 000 000 iterations is performed for all the ΔV_T shown in Fig. 3 for different temperatures. The result is the grid in Fig. 4 which shows where the retention is stable or not.

In contrast to the retention, for a functional read operation, the PG has to be weaker than the PU. Fig. 1(b) presents the critical path of the read operation. A read fail is caused by the difference of currents driven by the PG and PU but also by the difference between the capacitance of the bitline (Cblt) and the internal node (Cblti). When the bitcell is read, the voltage of the internal node at V_{DD} will drop, and if this drop is critical, the bitcell will inevitably toggle (leading to the data loss). The critical path of the read operation can be modeled as an *RC* network by replacing the nMOS and

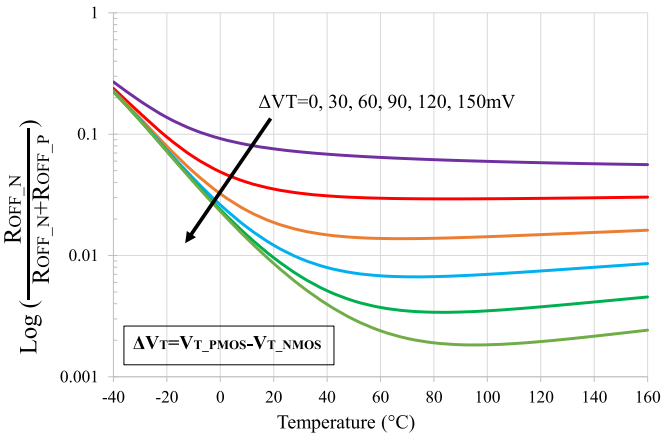

Fig. 3. Resistance ratio of voltage divider versus temperature for several ΔV_T (V_{T_pMOS}–V_{T_nMOS}).

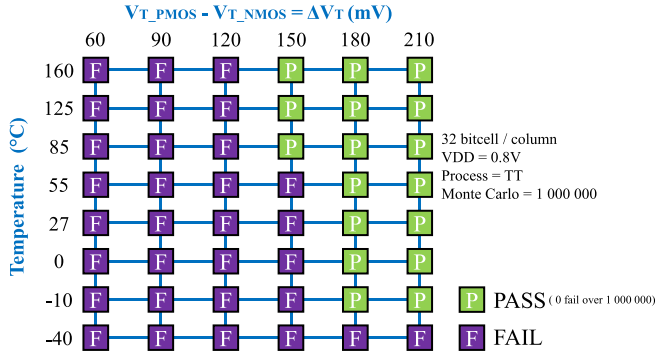

Fig. 4. Retention stability grid for several ΔV_T and temperatures.

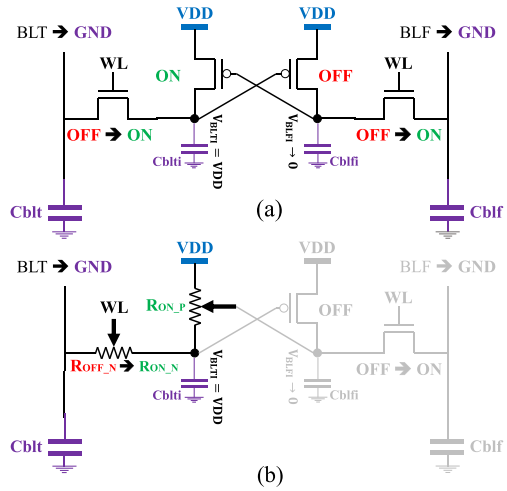

Fig. 5. (a) MOS based and (b) simplified model of the critical path during a read operation of 4T DL SRAM bitcell.

pMOS by their equivalent resistance and adding Cblti and Cblt as shown in Fig. 5.

To study the effect of Cblt and $R_{\mathrm{ON\_}P}$ on the voltage drop of the internal node during a read operation, we consider Cblti and $R_{\mathrm{ON\_}N}$ values to be constant. First, we focus on Cblt which is induced by the number of b/c. Fig. 6(a) shows the voltage drop of the internal node during a read operation for different number of b/c. We can see that the greater the number of b/c is, the greater the time of convergence of the internal node to V_{DD}. The intensity of the voltage drop is barely affected by the value of the Cblt. Second, we observe the effect of

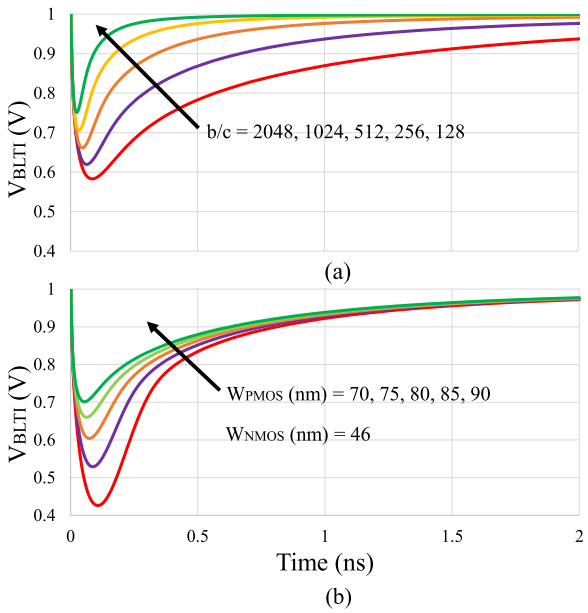

Fig. 6. Internal node voltage drop for (a) several b/c and (b) pMOS widths.

varying the pMOS width ($R_{\mathrm{ON\_}P}$) on the voltage drop of the internal node. Fig. 6(b) shows the voltage drop of the internal node during a read operation for different value of the pMOS width. We can see that the smaller the pMOS width is (greater $R_{\mathrm{ON\_}P}$), the greater the internal voltage drop will be.

From those results and observations, we distinguish two read fail mechanisms.

1) *A Fast Fail:* where the voltage drop of the internal node at V_{DD} is critical and leads to a bitcell toggle. This is driven by the ratio of the width between the nMOS and pMOS.
2) *A Slow Fail:* where the voltage drop is not critical but the voltage of the complementary node rises from G_{ND} toward V_{DD} and eventually provokes a bitcell toggle. This is due to the Cblt induced by the number of b/c.

The internal voltage drop is given by

$$\Delta V_{\mathrm{BLTI}} = \frac{R_{\mathrm{ON\_}N}}{R_{\mathrm{ON\_}N} + R_{\mathrm{ON\_}P}} \times V_{\mathrm{DD}}. \tag{3}$$

Hence, if we want the internal node voltage drop to be moderate, we must have

$$R_{\mathrm{ON\_}P} < R_{\mathrm{ON\_}N} \quad (4) \quad \rightarrow \quad \boxed{W_{\mathrm{pMOS}} > W_{\mathrm{nMOS}}}. \tag{5}$$

Fig. 7 presents internal-node voltage evolution under a read operation performed with different temperatures ranging from −40 °C to 160 °C. We can notice that the drop of the internal node becomes more important with the rise of temperature. For 140 °C and 160 °C, we can see that the drop is critical and leads to the bitcell toggle. It has to be noted that the rise of the internal node BLFI voltage is a consequence of the voltage drop of BLTI. With the rise in temperature, the V_T of both nMOS and pMOS decreases (as expected) but not at the same rate. Fig. 8(a) presents the impact of the temperature on the V_T ratio of nMOS and pMOS. The V_TnMOS decreases faster than V_TpMOS, which means that the current driven by the nMOS will become higher than one of the pMOS, which degrades the stability of the bitcell during a read operation.

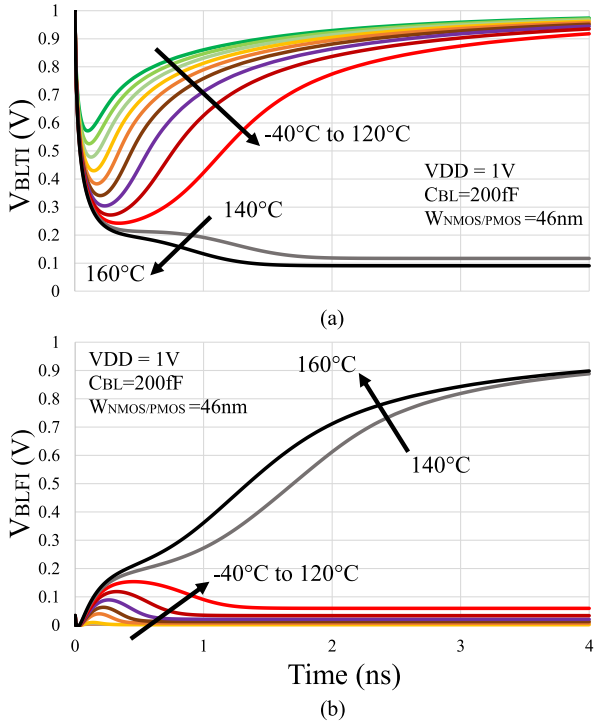

Fig. 7. Temperature effect on stability of the internal nodes (a) BLTI and (b) BLFI during a read operation.

Fig. 8(b) presents the impact of the temperature on the ON-state resistance of nMOS and pMOS. As for the V_T ratio, the ratio of the ON-state resistances, which defines the voltage drop of the internal node, decreases when temperature rises leading to a larger voltage drop.

Fig. 8(c) presents the impact of the temperature on the ratio of the ON-state currents of nMOS and pMOS, which is less favorable at high temperature since the nMOS becomes stronger than the pMOS. In fact, the changes in the ratios of V_T and ON-state resistance have an impact together on the I_{ON} current of nMOS and pMOS.

III. Physical Considerations in 3-D CoolCube Technology

A 0.078-μm^2 6T SRAM bitcell is first designed as a reference bitcell using the planar LETI 14-nm FD-SOI technology. Second, 3-D 6T and 4T bitcells are designed using a design kit extension for LETI 14-nm FD-SOI technology based on CoolCube process. In this process, an nMOS tier is fabricated over a pMOS tier and intertier vias (35 nm in diameter and 85 nm in pitch) are used to make connections between the bottom and the top tiers.

Fig. 9 presents the 3-D 4T bitcell, and we can see the pMOS PU on the bottom layer with a width of 90 nm. Note the metal intermediate and the 3-D vias that connect the bottom and top layers. The top layers contain the nMOS PG (width of 46 nm) connected to the word line (WL) and the bitlines (BLT and BLF). The 3-D 4T bitcell is 405 nm long and 135 nm large, which yields an area of the 0.054 μm^2. The height or thickness of the bitcell is equal to 150 nm. Previous works [18], [19] present a 6T design exploration using 3-D sequential technology and demonstrate high performance and 30% area gain compared to planar 6T bitcell. However,

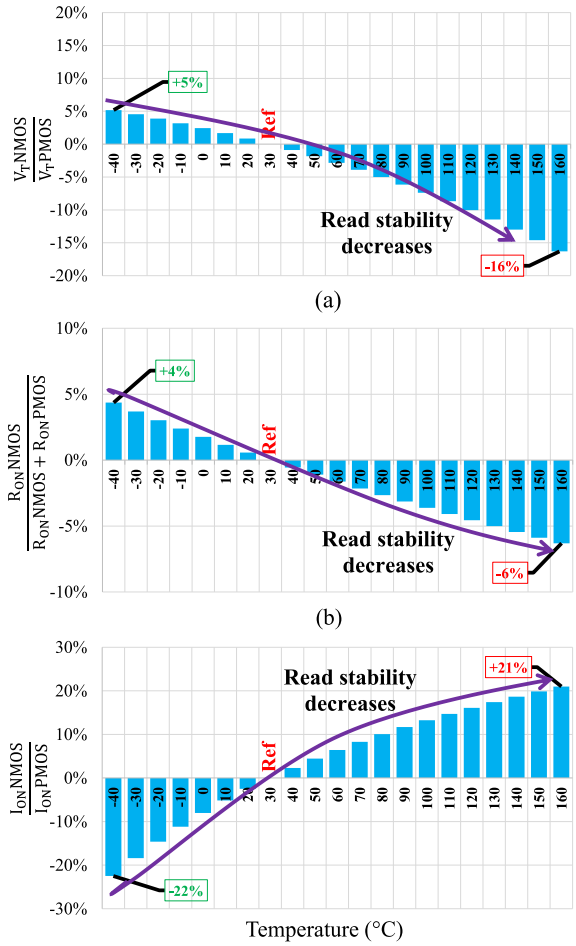

Fig. 8. Temperature effect on (a) V_T, (b) ON-state resistance, and (c) ON-state current of the nMOS and pMOS.

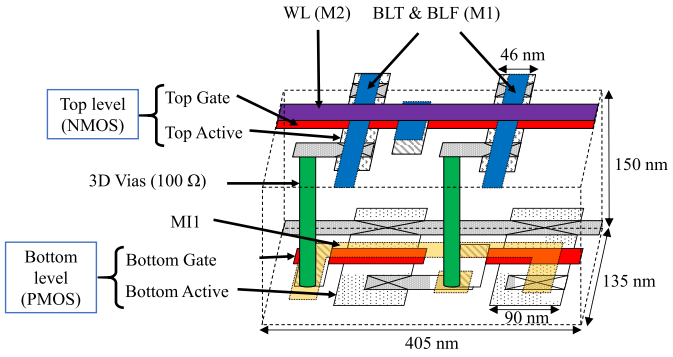

Fig. 9. 3-D picture of the 0.054 μm^2 4T DL SRAM bitcell in 3-D CoolCube technology [16].

the proposed layouts with intertier via just under top gate and active zone are currently not feasible in process fabrication. Thus, in practice, the area gain designing 6T bitcell in 3-D sequential technology is less than 5%. The 4T bitcell architecture achieves an area reduction of 30% compared to planar 6T bitcell (0.054 versus 0.078 μm^2).

IV. In-House Testbench and Simulation Results

The testbench setup for the study of the 4T DL SRAM bitcell is done in such a way it tests every operation and condition the bitcell could undergo.

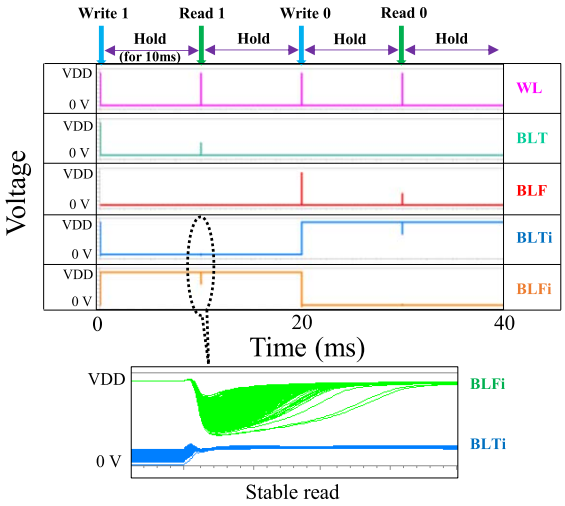

Fig. 10. Chronogram of the in-house testbench.

TABLE I
4T SRAM BITCELL V_T AND MOS SIZING CONFIGURATION

	W/L (nm)	NominalV_T (mV)	V_{BB}	FinalV_T (mV)	ΔV_T (mV) (V_{TPMOS} - V_{TNMOS})
NMOS	46/30	361	GND	361	179
PMOS	90/30	475	VDD	540	

Fig. 10 shows the global view of a full test. As we can see, we have mainly four sequences applied on the bitcell: hold-after-write "1," hold-after-read "1," hold-after-write "0," and hold-after-read "0." Ten milliseconds of retention time is set between each operation. This lapse of time is reasonably long for an SRAM bitcell, if a hold fail has to occur, it will happen as soon as the bitcell undergo a hold mode. The bitlines are precharged to G_{ND} instead of V_{DD} as in 6T bitcell because of the absence of pull-down nMOS. Therefore, bitlines are charged instead of being discharged during a read operation.

We performed the parasitic extraction of the 4T 3-D bitcell. The flow is based on Mentor Graphics Calibre X-ACT 3-D tool, in which several adjustments have been made to describe the 3-D sequential materials and architecture. Then, using the postlayout 4T bitcell netlist which includes the resistance of the 3-D vias connections, a sensitive analysis is performed with statistical simulations including global and local variations on a set of 1 000 000 MC simulations. Several design parameters are adjusted to optimize the bitcell robustness and performances.

The retention is investigated in Fig. 11, which shows the bitcell integrity during a retention phase for a temperature range between −10 °C and 125 °C and for supply voltages of 0.8 and 1 V. The fail percentage versus the V_T gap between the pMOS and nMOS are shown. Note that we have a robust retention with a 180-mV gap for $V_{DD} = 0.8$ V and 210 mV for $V_{DD} = 1$ V. These V_T gaps are obtained thanks to knobs (process and design) offered by the FD-SOI technology: the gate type combined with the backplane type [20] and back-biasing (BB) technique [21].

From those results, we decided to set the configuration shown in Table I for the 4T bitcell to guarantee a strong

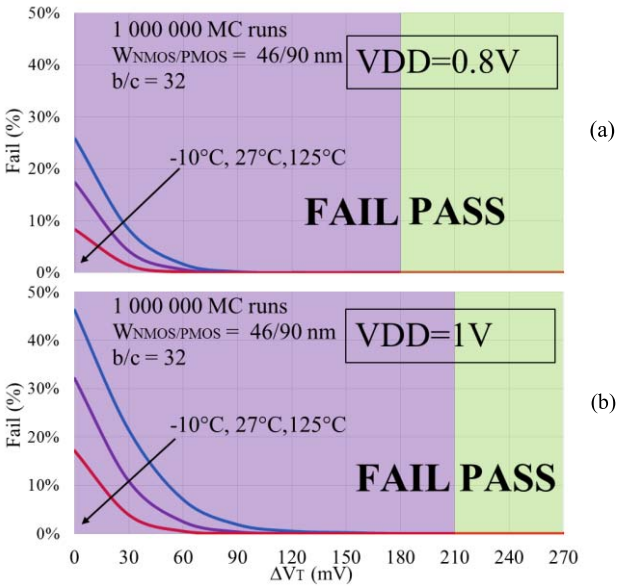

Fig. 11. Retention fail percentage versus the V_T gap at (a) $V_{DD} = 0.8$ V and (b) $V_{DD} = 1$ V.

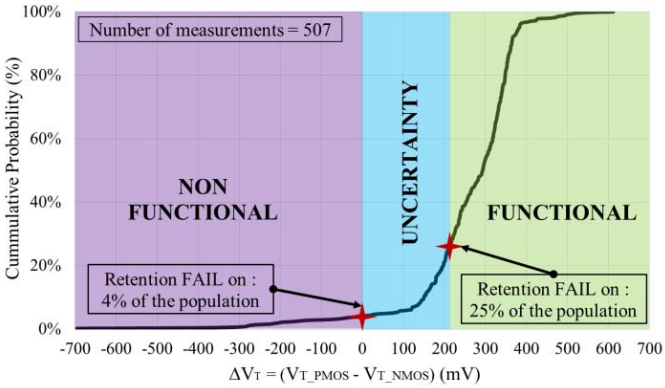

Fig. 12. Data retention validation with silicon data of V_T gap in LETI 14-nm planar FD-SOI technology.

retention mode. The widths of the nMOS and pMOS are the smallest possible to achieve a maximum density gain. The length (maximum enabled w/ CPP = 90 nm) is set at 30 nm ($L_{min} = 20$ nm) for both pMOS and nMOS in order to maintain a moderate variability.

The feasibility of this V_T gap has been validated with silicon data. A memory array of 1024 4T DL bitcells was fabricated on silicon with the 14-nm planar FD-SOI technology. These results can be extrapolated in 3-D technology since the V_T knobs are similar.

Eight wafers were processed to extract the global variation. On each wafers, 66 sites (dies) containing 1024 bitcells were measured to extract the local variation. On each site, four bitcells were selected and their pMOS and nMOS V_T measured.

Fig. 12 shows the cumulative probability of the silicon measurements of the V_T gap between the pMOS and the nMOS. We can observe that the configuration set in Table I is achievable at an industrial level. We obtain a fail in retention on a population ranging from 25% to 4%. The first value is

TABLE II
PLANAR AND 3-D V_T CONFIGURATION

	Planar		3D	
	NMOS	PMOS	TOP NMOS	BOT PMOS
Gate type	N	N	N	N
Back plane type	N	N	N	N
Well type	N	N	/	N
Mean V_T @VDD=1V	346 mV	608 mV	361 mV	570 mV
Retention OK on ##% of the population	75% - 96% (507 samples)		100% (10^6 samples)	

the worst case where we took only the cells that have a V_T gap $\geq$210 mV, and it is the V_T gap necessary for a functional retention at $V_{DD} = 1$ V in simulation. The second represents the best case, and this time we took the cells with a V_T gap $\geq$0 mV. We expect the real percentage to be between these two references since there will be functional bitcells with a V_T gap lower than 210 mV. Considering that these silicon measurements correspond to the first process in 14-nm FD-SOI for the 4T bitcell, the results are relatively acceptable.

Table II shows the V_T configuration used in the planar design (silicon measurements) and the 3-D design (simulation models). In both technologies, the V_T configuration is similar for the pMOS (n-BP on n-well). In the 3-D design, the nMOS is on the top layer and hence they lack a well layer [22]. But the presence or the type of the well does not affect the MOS V_T. Thus, the results validated in planar silicon are also valid in 3-D design.

Comparing the results of the planar design (silicon data) to the 3-D design (simulation data), we can notice that we have a lower percentage of population with functional retention in planar design than in 3-D design, even though we have a greater V_T gap in the planar measurements. The reason is related to the variability of the planar design silicon measurements which is greater than one of the 3-D design simulations. This is mainly due to the immaturity of the process. Through process optimization, we can gain maturity and reach a percentage close to 100%.

Using the configuration given in Table I, we investigated the 4T bitcell read stability. As mentioned previously, reading the bitcell represents a sensible operation since there is a risk to lose the data. A simple and straightforward solution is to strengthen the pMOS by enlarging its width as given in (5).

Fig. 13 shows the evaluation in read fail percentage for different number of b/c versus the pMOS width for the worst-case temperature (125 °C). We also display in Fig. 14 the necessary pMOS width (for no read fails) for a given number of b/c and the density gain compared to the planar 6T on the right axis. For example, in Fig. 14, for a number of 64 b/c, the minimum pMOS width is $W_{pMOS} = 120$ nm for a fully functional read operation, which yields a gain of 20% in density with respect to the planar 6T bitcell (0.078 μm^2). The maximum density gain is 30% and is possible for 32 b/c.

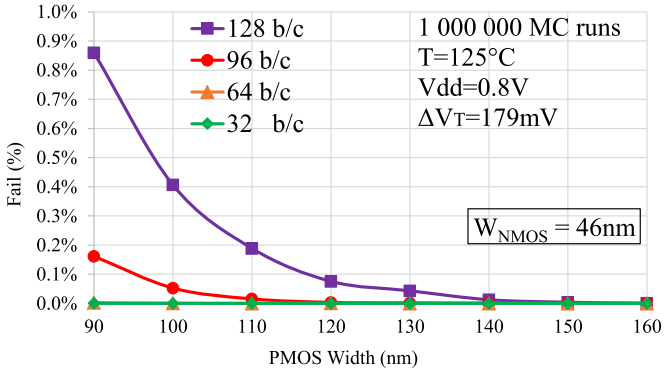

Fig. 13. Read fail percentage versus the pMOS width for several number of b/c.

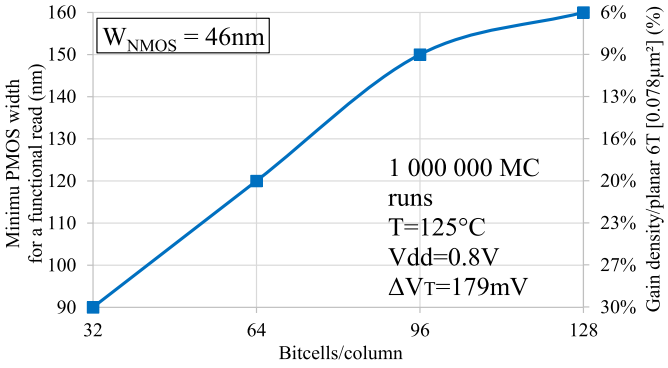

Fig. 14. Minimum pMOS width and gain density with respect to the planar 6T SRAM bitcell (0.078 μm^2) versus the number of b/c.

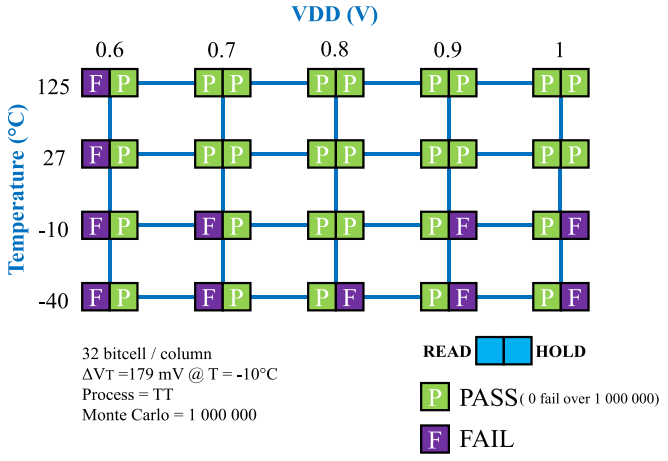

Fig. 15. PASS/FAIL grid for retention and read operations.

Fig. 15 presents PASS/FAIL grid for retention and read operations obtained after 1 000 000 MC simulations for the configuration of 32 bitcells/column with a pMOS width of 90 nm. We can see that the operating range, in which the bitcell is fully functional, covers 0.7, 0.8, 0.9, and 1 V of supply voltage, ranging from 27 °C to 125 °C. We can observe that the more the temperature decreases, the less the margin we have on the variation of V_{DD} supply.

V. 4T SRAM BITCELL ENHANCED WITH DYNAMIC BACK-BIASING TECHNIQUE

Dynamic back biasing (DBB) is an efficient design technique to increase the bitcell stability and/or improve its

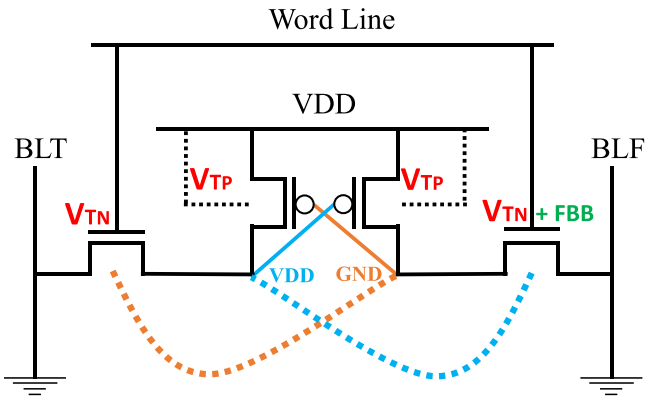

Fig. 16. 4T DL SRAM bitcell with data-dependent back biasing on top transistor (nMOS pass gate).

TABLE III

4T SRAM BITCELL V_T AND MOS SIZING IN DBB CONFIGURATION FOR THE (a) RETENTION CRITICAL PATH AND (b) READ CRITICAL PATH

(a)	W/L (nm)	Nominal V_T (mV)	V_{BB}	Final V_T (mV)	ΔV_T (mV) (V_{TPMOS} - V_{TNMOS})
NMOS	46/30	361	VDD	291	179
PMOS	90/30	405	VDD	470	

(b)	W/L (nm)	Nominal V_T (mV)	V_{BB}	Final V_T (mV)	ΔV_T (mV) (V_{TPMOS} - V_{TNMOS})
NMOS	46/30	361	GND	361	109
PMOS	90/30	405	VDD	470	

performance [23]. FD-SOI technology is intrinsically more suitable to exploit the full potential of DBB compared to others technologies [24] (e.g., Bulk, FinFET). In this section, we demonstrate how a data-dependent DBB can strengthen the stability during a read operation.

Fig. 16 presents DBB scheme for the 4T DL SRAM bitcell. The back gate of the PG nMOS is connected to the opposite internal node (BLTI or BLFI). This configuration allows the bitcell to maintain a strong retention state where the "0" is stored (BLFI) and stabilizes the reading by strengthening the PU pMOS connected to the stored "1" (BLTI).

The main challenge of this DBB configuration is to back bias each nMOS independently. Hence, the feasibility of this configuration has been investigated. We have developed a solution that does not affect the bitcell area by exploiting the 3-D CoolCube technology process flexibility. It is not presented here since it is still under investigation and will be the object of future publications.

Table III(a) and (b) shows, respectively, new V_T configurations that can be set to maximize the benefits of the DBB during the read and hold operations separately.

The DBB configuration presented in this section was set as follows: by applying DBB, we enhanced the retention robustness by enlarging the V_T gap. Hence, we lowered the V_T of the pMOS (in simulations) until we reached the V_T gap necessary for a good retention (179 mV).

Fig. 17 presents the necessary pMOS width (for no read fails) for a given number of b/c and the density gain compared to the planar 6T on the right axis for the configuration w/ and

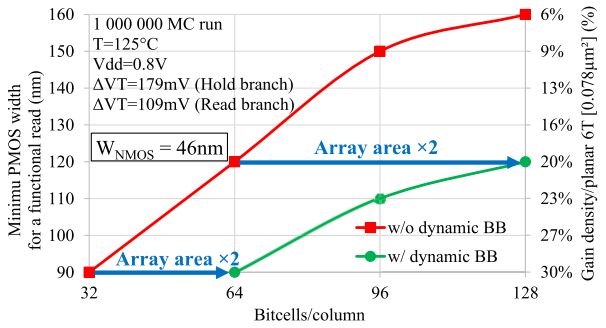

Fig. 17. Minimum pMOS width and gain density with respect to the planar 6T SRAM bitcell (0.078 μm^2) versus the number of b/c with and without DBB.

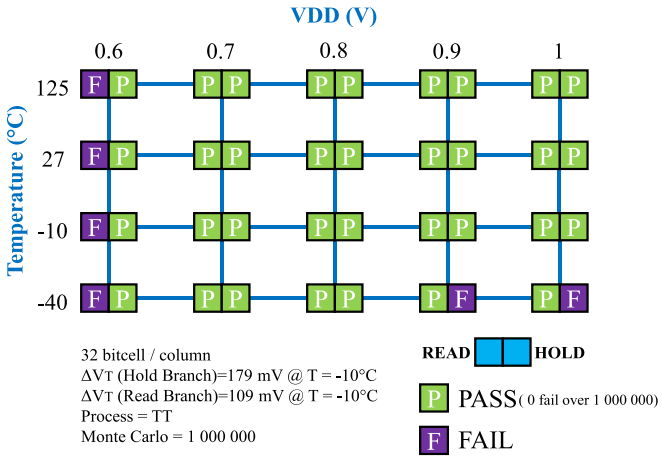

Fig. 18. PASS/FAIL grid for retention and read operations with DBB.

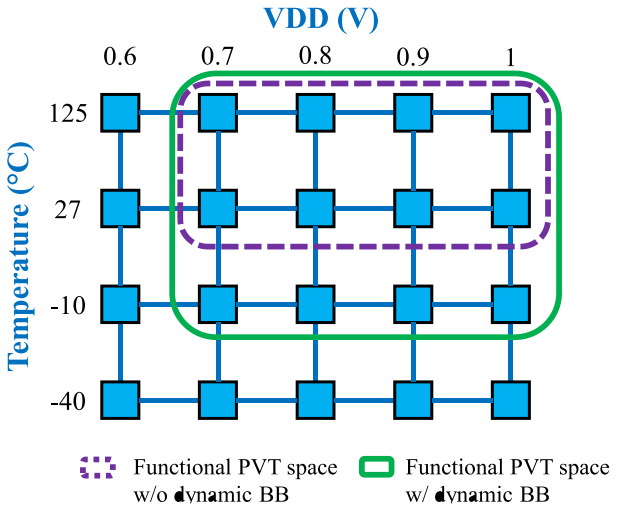

Fig. 19. PVT functional space w/ and w/o DBB.

w/o DBB. We can see a significant impact of the DBB on the ability of doubling the size of the memory array without enlarging the pMOS width.

Fig. 18 presents PASS/FAIL grid for retention and read operations obtained after 1 000 000 MC simulations for the configuration of 32 bitcells/column with a pMOS width of 90 nm and using the data-dependent DBB. Fig. 19 compares the results shown in Figs. 15 and 18, and we observe that the

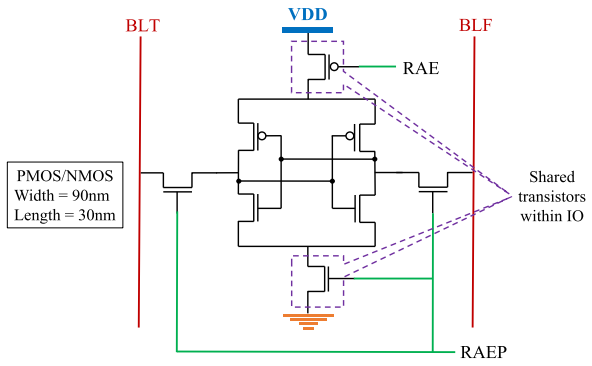

Fig. 20. Proposed RA schematic for the 4T DL SRAM bitcell.

DBB enlarged the operating voltage and temperature range. With the DBB configuration, the bitcell is considered as fully functional for a temperature ranging from −10 °C to 125 °C and a V_{DD} variation from 0.7 to 1 V.

VI. Proposed Read-Assist Technique

The main purpose of the design of the 4T bitcell is the density gain. With the 3-D design we showed a density gain on the bitcell level, in the previous section, we showed how a DBB could improve the density gain on the memory array level by stacking more b/c. In this section, we want to go further in the density gain on the memory array level.

Therefore, we present a RA circuit designed specifically for the 4T SRAM bitcell. This circuit is not new, and it is used nowadays as a write-back assist. Hence, it senses and writes back the data sensed [25]. In our case we use it differently, its purpose is to help obtain a sufficient voltage difference between the bitlines during the read operation. This RA circuit, shown in Fig. 20, works as follows: after the start of read operation, the RA senses the bitlines to know the content of the selected bitcell. Then, it writes back consequently the bitlines in the right way (same way the bitcell does) to support the bitcell and complete the read operation by stabilizing the bitcell content.

This RA circuit is placed at the bottom of each column between the bitcell array and the conventional I/O as shown in Fig. 21, and fits into the array column width. Its role is to assist the bitcell during a read operation and not to perform the read itself. Hence, there is the necessity to add a conventional sense amplifier [26] for a proper read. The dimension of RA, designed in 3-D CoolCube technology, is evaluated as equivalent to five times the bitcell height, which represents 4% of the area of a column containing 128 bitcells.

Note that for this configuration, we have chosen an interleaving structure with a 4 to 1 multiplexer. This can be modified as wanted. There is no risk to lose the data on the column that are not selected since they will be stressed by a read operation. And the stability of this operation has been validated.

Thanks to this assist, we can align 128 bitcell in a column without having to enlarge the pMOS (PU) width. Thus, this RA circuit allows us to achieve the maximum density possible for the 4T bitcell. Fig. 22 presents the necessary pMOS

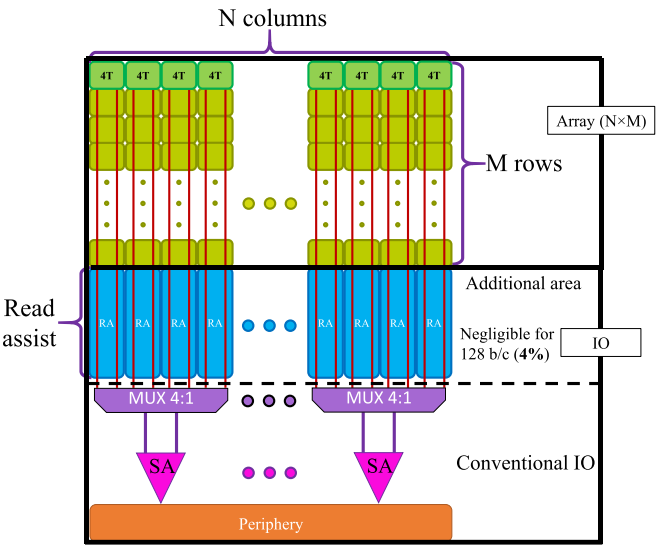

Fig. 21. Placement of the RA on the memory array.

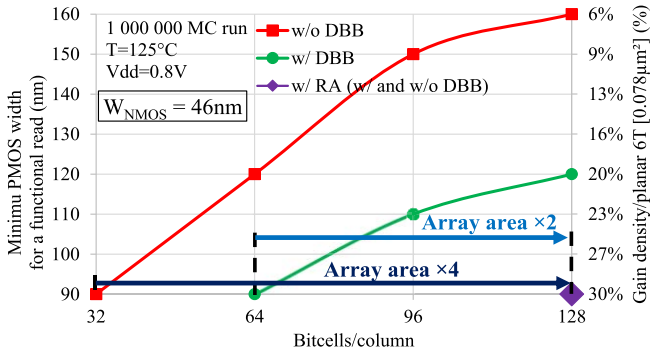

Fig. 22. Minimum pMOS width and gain density with respect to the planar 6T SRAM bitcell (0.078 μm^2) versus the number of b/c with and without DBB and with the read assist.

width (for no read failures) for a given number of b/c and the density gain compared to the planar 6T on the right axis for the configuration w/ and w/o DBB and with the addition of the RA. For 128 b/c, we can achieve a density gain of 6% (compared to the planar 6T) without DBB and 20% with DBB. With the proposed RA circuit, we can reach a density gain up to 30%. The 1 000 000 MC simulations are performed to determine the operating range of the 4T SRAM bitcell combined with the proposed RA. From the results obtained, we noticed that it is the same w/ or w/o the RA, whether it is a config'uration w/o DBB (Fig. 15) or w/ DBB (Fig. 18). Therefore, results are not shown in this paper.

The drawback of this design is that it slows down the read operation. But, since the objective of the 4T bitcell is to be as dense as possible (and not as fast as possible), this solution remains adequate.

Fig. 23 presents the chronogram of the read operation with the proposed RA circuit.

To avoid fails, we read the bitcell in three main phases.

1) *Phase 1:* It consists of setting the RA circuit by precharging its internal nodes. We simply pull the internal node at G_{ND} by activating (high) the RAEP signal.

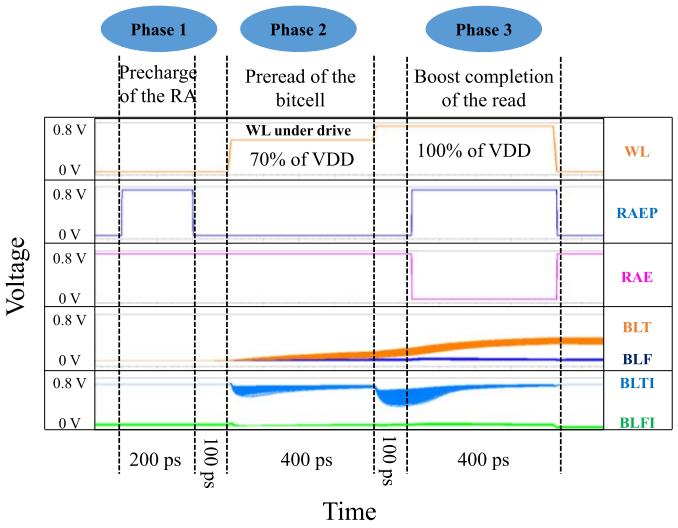

Fig. 23. Chronogram of the read operation with the proposed read assist.

TABLE IV

READ TIME, WRITE TIME, AND DENSITY GAIN W/ AND W/O DBB AND PROPOSED RA CONFIGURATION

4T SRAM bitcell configuration **128 b/c**	Read time (ps)	Write time (ps)	Gain density/planar 6T (0.078µm²)
Without **DBB** without **RA**	**65**	44	**6%**
With **DBB** without **RA**	**64**	37	**20%**
Without **DBB** with **RA**	**625**	44	**30%**
With **DBB** with **RA**	**345**	37	**30%**

2) *Phase 2:* In this phase, we charge the bitline to enable RA circuit to presense the data. But, to avoid any fail, we set the WL voltage at 70% of V_{DD} (WL under drive-assist technique [27], [28]).
3) After the end of phase 2, the WL voltage is set to 100% of V_{DD} to fully activate the PG. After 100 ps, it will create enough voltage difference (100 mV) between the bitlines for the RA to work properly. If the RA is not activated under 100 ps after that the read operation could fail.
4) *Phase 3:* The read operation is completed in this phase. The RA is activated; it senses the data stored in the bitcell and pulls each bitline in the same way (up or down) the bitcell pulled them in the precedent phases.

The three-phase scheme enables fail-free read operation with a minimum of 100-mV bitlines voltage difference reached for a proper read.

Table IV summarizes the difference between the possible configurations with the DBB and the RA. It presents for each configuration, the read and write operation time, and density gain compared to the planar 6T.

From the results, we can see the efficiency of the DBB in stabilizing the read operation, thus allowing a greater density gain. The RA can be used in the case when the time of the read operation is not *a priori*ty but density gain is.

The application targeted by our SRAM memory would depend on the use of the presented RA. If not used, then the SRAM will possess good performances with fast read and write operations, but the matrix size would be limited. In that case, L1 or L2 cache level would be suitable.

If the RA is used, bigger matrices would be achievable but at the expense of speed (read time). In that case, L3 cache level would be the appropriate application.

VII. CONCLUSION

This paper presents a 4T SRAM bitcell designed using the 3-D CoolCube technology. A density gain up to 30% is achieved compared to a planar 6T bitcell (0.078 μm^2) in the same technology node. Based on an in-house simulation testbench, the failure mechanisms of the 4T bitcell have been deeply studied and suitable solutions are proposed to overcome the stability issues. Using the FD-SOI technology and the 3-D CoolCube technology, we have demonstrated the efficiency of a DBB as a design knob to considerably improve the stability of the 4T bitcell. Moreover, thanks to the DBB, the size of our array can be double by doubling the number of bitcells per column. We also presented a RA technique that stabilizes the reading operation at the expenses of the read time, but allow to increase considerably the number of bitcell per column. Finally, a considerable number of bitcell per column (up to 128) with 30% density gain compared to the planar 6T is demonstrated.

REFERENCES

[1] B. Davari, R. H. Dennard, and G. G. Shahidi, "CMOS scaling for high performance and low power-the next ten years," *Proc. IEEE*, vol. 83, no. 4, pp. 595–606, Apr. 1995.
[2] T. Skotnicki, J. A. Hutchby, T.-J. King, H. S. P. Wong, and F. Boeuf, "The end of CMOS scaling: Toward the introduction of new materials and structural changes to improve MOSFET performance," *IEEE Circuits Devices Mag.*, vol. 21, no. 1, pp. 16–26, Jan. 2005.
[3] A. Gill, C. Madhu, and P. Kaur, "Investigation of short channel effects in Bulk MOSFET and SOI FinFET at 20 nm node technology," in *Proc. Annu. IEEE India Conf. (INDICON)*, Dec. 2015, pp. 1–4.
[4] S. K. Saha, "Modeling process variability in scaled CMOS technology," *IEEE Design Test Comput.*, vol. 27, no. 2, pp. 8–16, Mar./Apr. 2010.
[5] Y.-S. Lin *et al.*, "Leakage scaling in deep submicron CMOS for SoC," *IEEE Trans. Electron Devices*, vol. 49, no. 6, pp. 1034–1041, Jun. 2002.
[6] D. Maji *et al.*, "A junction leakage mechanism and its effects on advance SRAM failure," in *Proc. IEEE Int. Rel. Phys. Symp. (IRPS)*, Apr. 2013, pp. 3E.1.1–3E.1.5.
[7] K. Athikulwongse, D. H. Kim, M. Jung, and S. K. Lim, "Block-level designs of die-to-wafer bonded 3D ICs and their design quality tradeoffs," in *Proc. Design Autom. Conf. (ASP-DAC)*, 2013, pp. 687–692.
[8] K. Puttaswamy and G. H. Loh, "3D-integrated SRAM components for high-performance microprocessors," *IEEE Trans. Comput.*, vol. 58, no. 10, pp. 1369–1381, Oct. 2009.
[9] L. Brunet *et al.*, "First demonstration of a CMOS over CMOS 3D VLSI CoolCube integration on 300 mm wafers," in *Proc. IEEE Symp. VLSI Technol. (VLSIT)*, Jun. 2016, pp. 1–2.
[10] P. Batude *et al.*, "3DVLSI with CoolCube process: An alternative path to scaling," in *Proc. Symp. VLSI Technol. (VLSIT)*, 2015, pp. T48–T49.
[11] F. Clermidy, O. Billoint, H. Sarhan, and S. Thuries, "Technology scaling: The CoolCube paradigm," in *Proc. IEEE SOI-3D-Subthreshold Microelectron. Technol. Unified Conf. (SS)*, Oct. 2015, pp. 1–4.

[12] J.-P. Noel, O. Thomas, C. Fenouillet-Beranger, M.-A. Jaud, and A. Amara, "Robust multi-VT 4 T SRAM cell in 45 nm thin BOx fully-depleted SOI technology with ground plane," in *Proc. IEEE Int. Conf. IC Design Technol.*, May 2009, pp. 191–194.
[13] A. Shafaei and M. Pedram, "Energy-efficient cache memories using a dual-Vt 4 T SRAM cell with read-assist techniques," in *Proc. Design, Autom. Test Eur. Conf. Exhibit. (DATE)*, Mar. 2016, pp. 457–462.
[14] C. Lage, J. D. Hayden, and C. Subramanian, "Advanced SRAM technology-the race between 4 T and 6 T cells," in *Proc. Int. Electron Devices Meeting*, Dec. 1996, pp. 271–274.
[15] J. Hartmann, "FD-SOI technology development and key devices characteristics for fast, power efficient, low voltage SoCs," in *Proc. IEEE Compound Semiconductor Integr. Circuit Symp. (CSICS)*, Oct. 2014, pp. 1–4.
[16] M. Brocard *et al.*, "High density SRAM bitcell architecture in 3D sequential CoolCube 14 nm technology," in *Proc. IEEE SOI-3D-Subthreshold Microelectron. Technol. Unified Conf. (SS)*, Oct. 2016, pp. 1–3.
[17] D. Yagain, A. Parakh, A. Kedia, and G. K. Gupta, "Design and implementation of high speed, low area multiported loadless 4 T memory cell," in *Proc. 4th Int. Conf. Emerg. Trends Eng. Technol.*, 2011, pp. 268–273.
[18] C. Liu and S. K. Lim, "Ultra-high density 3D SRAM cell designs for monolithic 3D integration," in *Proc. IEEE Int. Interconnect Technol. Conf.*, Jun. 2012, pp. 1–3.
[19] O. Thomas, M. Vinet, O. Rozeau, P. Batude, and A. Valentian, "Compact 6 T SRAM cell with robust read/write stabilizing design in 45 nm Monolithic 3D IC technology," in *Proc. IEEE Int. Conf. IC Design Technol.*, May 2009, pp. 195–198.
[20] J.-P. Noel, O. Thomas, C. Fenouillet-Beranger, M.-A. Jaud, P. Scheiblin, and A. Amara, "A simple and efficient concept for setting up multi-VT devices in thin BOx fully-depleted SOI technology," in *Proc. Eur. Solid State Device Res. Conf.*, Sep. 2009, pp. 137–140.
[21] V. Asthana, M. Kar, J. Jimenez, J.-P. Noel, S. Haendler, and P. Galy, "Circuit optimization of 4 T, 6 T, 8 T, 10 T SRAM bitcells in 28 nm UTBB FD-SOI technology using back-gate bias control," in *Proc. Eur. Solid State Device Res. Conf. (ESSDERC)*, 2013, pp. 415–418.
[22] P. Batude *et al.*, "3D monolithic integration," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2011, pp. 2233–2236.
[23] B. Ebrahimi, A. Afzali-Kusha, and N. Sehatbakhsh, "Robust polysilicon gate FinFET SRAM design using dynamic back-gate bias," in *Proc. Design Technol. Integr. Syst. Nanosc. Era (DTIS)*, 2013, pp. 171–172.
[24] Q. Liu *et al.*, "Impact of back bias on ultra-thin body and BOX (UTBB) devices," in *Proc. Symp. VLSI Technol. (VLSIT)*, Jun. 2011, pp. 160–161.
[25] J.-J. Wu *et al.*, "A large σV_{TH}/VDD tolerant zigzag 8 T SRAM with area-efficient decoupled differential sensing and fast write-back scheme," *IEEE J. Solid-State Circuits*, vol. 46, no. 4, pp. 815–827, Apr. 2011.
[26] Y.-P. Tao and W.-P. Hu, "Design of sense amplifier in the high speed SRAM," in *Proc. Cyber-Enabled Distrib. Comput. Knowl. Discovery (CyberC)*, 2015, pp. 384–387.
[27] V. P.-H. Hu, M.-L. Fan, P. Su, and C.-T. Chuang, "Analysis of GeOI FinFET 6 T SRAM cells with variation-tolerant WLUD read-assist and TVC write-assist," *IEEE Trans. Electron Devices*, vol. 62, no. 6, pp. 1710–1715, Jun. 2015.
[28] A. Kumar, G. S. Visweswaran, V. Kumar, and K. Saha, "A 0.5 V VMIN 6 T SRAM in 28 nm UTBB FDSOI technology using compensated WLUD scheme with zero performance loss," in *Proc. 29th Int. Conf. VLSI Design 15th Int. Conf. Embedded Syst. (VLSID)*, Jan. 2016, pp. 191–195,

Réda Boumchedda was born in Lyon, France, in 1991. He received the B.S. degree in electronic and electrical engineering from the Institute of Electrical and Electronic Engineering, University of Boumerdès, Boumerdès, Algeria, in 2013, and the M.S. degree in microelectronics from Grenoble University, Grenoble, France, in 2015. He is currently pursuing the Ph.D. degree with STMicroelectronics, Crolles, France, and Grenoble University, in collaboration with the memory design team of the CEA-LETI Research Center, Grenoble.

His current research interests include the application of dynamic back bias in memories.

Jean-Philippe Noel was born in Alès, France, in 1985. He received the M.S. degree in microelectronics from Polytech Marseille, Marseille, France, in 2008, and the Ph.D. degree in microelectronics from the University of Grenoble, Grenoble, France, in 2011.

He was with the CEA-LETI, Grenoble, where he was involved in the design optimization of low-power digital and memory circuits in ultra thin body & box fully depleted silicon on insulator (UTBB FD-SOI) technology. In 2011, he joined STMicroelectronics, Crolles, France, as a Memory Design Engineer. He was involved in the design of innovative SRAM and ternary content addressable memory circuits in UTBB FD-SOI technology. Since 2016, he has been with CEA-LETI, focusing on the circuit design of innovative SRAM/TCAM and emerging nonvolatile memories, such as ReRAM and PCRAM.

Bastien Giraud received the Ph.D. degree from Telecom ParisTech, Paris, France, in 2008. His Ph.D. thesis focused on SRAM design in Double-Gate FD-SOI.

In 2009, he joined the University of California, Berkeley, CA, USA, as a Post-Doctoral Researcher, where he was involved in low-power circuits and SRAM variability. Since 2010, he has been a Circuit Designer with the CEA-LETI, Grenoble, France, where he is involved in memory and low-power circuit in advanced technologies. He has authored more than 25 papers in international conferences and journals. He has authored a book chapter and is the main Inventor or Co-Inventor of 15 patents. His current research interests include resilient memory with assist technics, energy efficiency, specific design technics, and nonvolatile memories, and also include SRAM ultra low voltage and robust, smart content addressable memory, logic in memory, crossbar, using advanced CMOS technologies and nonvolatile RRAM technologies such as CBRAM, OxRAM, and PCRAM.

Kaya Can Akyel was born in Adana, Turkey, in 1986. He received the B.Sc. and M.Sc. degrees in electrical and computer engineering from Joseph Fourier University, Saint-Martin-d'Hères, France, in 2009 and 2011, respectively, and the Ph.D. degree in micro- and nanoelectronics from the University of Grenoble Alpes, Grenoble, France, in 2014, with a focus on the development of SPICE-level simulation methodologies for modeling of the process variability impact on deep-submicrometer SRAMs.

In 2015, he joined the CEA-LETI Research Center, Grenoble, as a Post-Doctoral Researcher of the memory design team. His current research interests include innovative memory circuit design for novel computing methods and reconfigurable circuits.

Mélanie Brocard was born in Grenoble, France, in 1987. She received the Engineering degree in electronics and electrotechnics from the Grenoble National Polytechnical Institute, Grenoble, and the Ph.D. degree from the University of Grenoble, Grenoble, in 2013.

From 2010 to 2013, she was with STMicroelectronics, Crolles, France, in collaboration with IMEP-LAHC Chambery, Bourget-du-Lac, France, and CEA-LETI, Grenoble, on the radio frequency characterization of the electromagnetic coupling between device and through silicon via. Then, she was with the technology to device team, STMicroelectronics. She is currently a Designer with CEA-LETI, where she is involved in the research on 3-D monolithic technology. Her current research interests include device and coupling modeling, design of 3-D SRAM bitcell and standard cells, SPICE level characterization, and thermal characterization.

David Turgis received the master's degree in microelectronics from ENSICaen, Caen, France, in 1997.

Since 1999, he has been involved in memory compilers in CMOS and non volatile memories technologies. He is currently a Library Manager and a Senior Expert with STMicroelectronics, Crolles, France. He is the Memory Manager of the SRAM/BIST team in Crolles near Grenoble, France.

Edith Beigne joined CEA-LETI, Grenoble, France, in 1998. Since 2009, she has been a Senior Scientist with the Digital and Mixed-Signal Design Laboratory, where she researches low-power and adaptive circuit techniques, exploiting asynchronous design and advanced technology nodes, such as FDSOI 28 nm and 14 nm for many different applications from high-performance MPSoC to ultralow-power IoT applications. She has authored and co-authored over 100 publications.

Dr. Beigne has been with the ISSCC TPC since 2014 and the VLSI Symposium since 2015. She is a SSCS Distinguished Lecturer from 2016 to 2017.