A Self-Refreshable Bit-Cell for Single-Cycle Refreshing of Embedded Memories

Binyamin Frankel and Shmuel Wimer, *Member, IEEE*

Abstract—Power supply voltage reduction is a primary enabler for sustaining the increasing demand for ultra-low power processors. On-die memories, which are traditionally implemented by SRAM, stop functioning properly when the supply voltage is scaled down aggressively; hence, embedded DRAM (eDRAM) bit-cells are used instead. These bit-cells leak their data strongly in one direction, whereas the leakage in the opposite direction is considerably lower. Due to their intrinsic limited Data Retention Time (DRT), these memories require power-hungry refreshing, which degrades performance. In an attempt to extend the DRT of a bit-cell, theoretically to infinity, compounds of various types of storage nodes in a single bit-cell, storing the datum and its complement, were examined here. A rigorous proof shows that under realistic leakage models, there is an inherent incompleteness preventing the proper readout and decision of the stored value after a certain time. Adopting the idea of dual-polarity complementary storage nodes, a new eDRAM self-refreshable bit-cell is proposed that yields a considerably extended DRT. The dual-polarity property enables the refreshing of an entire memory array in a single clock cycle, thus almost nullifying the unavoidable performance loss occurred by row-by-row ordinary power-hungry refreshing.

Index Terms—Memory design, memory technologies, dynamic memories, embedded memories, refreshing

1 Introduction

This paper proposes a new eDRAM CMOS compatible bit-cell, potentially overcoming the Data Retention Time (DRT) limitations and their refreshing overheads. In what follows we use the terms data and value interchangeably, as we also do for memory row and word.

Power supply voltage reduction is a primary enabler for sustaining the increasing demand for ultra-low power processors. Their on-die memories are traditionally implemented by 6T-SRAM which stops operating robustly when the supply voltage is scaled down aggressively [1]; hence, embedded DRAM (eDRAM) bit-cells are used instead. One type of eDRAM comprises conventional 1-transistor-1-capacitor (1T1C) bit-cells. Though affording very high density, its CMOS incompatibility requires special process technology which involves very high manufacturing costs. A Gain-Cell (GC) CMOS compatible solution was used in [2] for high-speed on-die caches. Though the GC bit-cell footprint is larger than 1T1C, its area is much smaller than the SRAM and consumes less power [3].

Dynamic memory cells are intrinsically subject to data leakage over time. Depending on the bit-cell polarity, both 1T1C and GC leak their data strongly in one direction; namely, $1 \rightarrow 0$ or $0 \rightarrow 1$, whereas the leakage in the opposite direction; namely, $0 \rightarrow 1$ or $1 \rightarrow 0$, respectively, is considerably lower. A bit-cell is called an N-type if it leaks $1 \rightarrow 0$ strongly and $0 \rightarrow 1$ weakly, and a P-type if it leaks $0 \rightarrow 1$ strongly and $1 \rightarrow 0$ weakly.

• *The authors are with Engineering Faculty, Bar-Ilan University, Ramat-Gan 52900, Israel. E-mail: binyamin.frankel@gmail.com, shmuel.wimer@biu.ac.il.*

Manuscript received 20 Feb. 2021; revised 14 Feb. 2022; accepted 6 Mar. 2022. Date of publication 10 Mar. 2022; date of current version 13 Jan. 2023.
This work was supported by Israel Innovation Authority (MAGNET program) under the GenPro Consortium.
(Corresponding author: Shmuel Wimer.)
Recommended for acceptance by U. R. Karpuzcu.
Digital Object Identifier no. 10.1109/TC.2022.3158481

Because of their limited DRT, these memories require power-hungry refreshing, which degrades performance. The solid curves in Fig. 1 present the typical leakage properties of dynamic memory N-type cells that leak $1 \rightarrow 0$ strongly and $0 \rightarrow 1$ weakly. The dashed curves in Fig. 1 present the typical leakage properties of dynamic memory P-type cells that leak $1 \rightarrow 0$ weakly and $0 \rightarrow 1$ strongly. Roughly speaking, the DRT is defined by the time elapsed since a value Z has been written into the cell until it is impossible to read it out and conclude whether it was written as $Z = 0$ or $Z = 1$. Depending on the underlying technology, the DRT can range from tens of micro-seconds in GCs to a few seconds in 1T1C [4]. Once the DRT has elapsed after writing, a refreshing to restore the initially stored value must take place. DRT elaboration and various refreshing algorithms are discussed in [5], [6], [7].

In an attempt to extend DRT, theoretically to infinity, this brief report examines two strategies. The first considers compounding within the same bit-cell storage nodes of both the N-type and the P-type, two of each type, that store the datum and its complement. It is predicated on the assumption that due to the differences in leakage characteristics of the N-type and the P-type storage nodes, the proper datum can always be read out and deduced correctly regardless of the time elapsed after it has been written. This is discussed in Section 3. The second strategy relies on a novel self-refreshable bit-cell circuit architecture, possessing symmetric leakage characteristics [8], as discussed in Sections 4 and 5. A brief description of the data leakage that occurred in GCs is presented first.

2 The Leakage of GC Bit-Cell

This section considers GC bit-cells, although the subsequent analysis and conclusions apply to other bit-cell types as well. Fig. 2 shows the simplest GC, which is comprised of two transistors, called a 2T-GC. This GC can be of the N-type [9] or the P-type [10], henceforth denoted by N as depicted in Fig. 2a, and P as depicted in Fig. 2b, respectively. The value written into the GC is stored at the SN node, which is made up of the gate capacitance of transistor MR and the source/drain capacitance of transistor MW.

Data are written into N-GC in Fig. 2a from the Write Bit-Line (WBL) by setting the Write Word-Line (WWL) to 1, thus turning the MW on. The data are then stored at the SN. The readout of the SN is done by first pre-charging the Read Bit-Line (RBL) to 1, and then setting the Read Word-Line (RWL) to 0. The data stored at the SN are then obtained at the RBL in the opposite polarity; in other words, if $\mathrm{SN} = 1$ then $\mathrm{RBL} = 0$ and vice-versa. The $\mathrm{P}-$GC in Fig. 2b works similarly but the voltage levels operate oppositely. Other combinations of 2T-GC transistors where the SN is read out positively are discussed in [11]. To avoid confusion, it is assumed in what follows that $\mathrm{RBL} = \mathrm{SN}$; namely, the readout is taken positively.

The values stored in the SN leak over time due to inherent leakage paths. This leakage, however, is asymmetric with respect to 0 and 1. N-GC leaks $1 \rightarrow 0$ strongly and $0 \rightarrow 1$ weakly. Subthreshold conduction (I_{sub}) and gate leakages (I_{gate}) exist for both $\mathrm{SN} = 0$ and $\mathrm{SN} = 1$. $\mathrm{SN} = 1$ nevertheless incurs junction leakage (I_{diff}), gate induced drain leakage (I_{GIDL}), and edge-direct tunneling leakage (I_{EDT}) which do not exist for $\mathrm{SN} = 0$. A similar situation for the opposite polarities exists in P-GC [12], [13].

The GC's DRT is limited because of the leakage mechanisms present in all MOS devices. To understand how data leak from the SN, assume that $\mathrm{SN} = 1$ was written. The SN is isolated from the WBL by the off MW (except at write). However, in the worst-case $\mathrm{WBL} = 0$ for the majority of the writes of other words. This

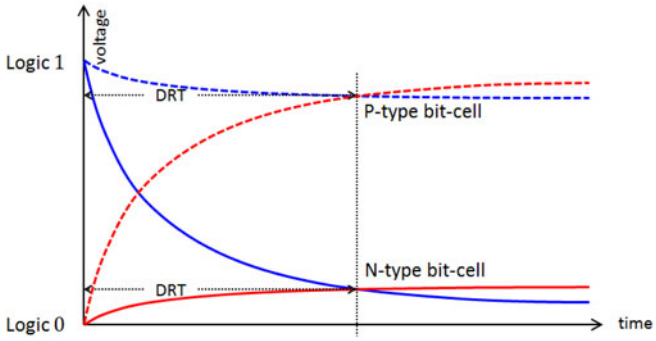

Fig. 1. Data retention time of a dynamic memory cell.

causes $\mathrm{SN}\ 1 \rightarrow 0$ leakage through the non-ideal switch MW. $\mathrm{SN}\ 1 \rightarrow 0$ leakage is shown in Fig. 3a by the blue curves obtained by Monte-Carlo SPICE simulations [9]. The opposite $0 \rightarrow 1$ leakage occurs when $\mathrm{SN} = 0$ and $\mathrm{WBL} = 1$, but as clearly shown by the red curves, it is slower by far.

More complex GCs aiming at longer DRT behave similarly, as shown in Fig. 3b for 5T-GC [14]. The P-GC in Fig. 3c depicts the mirror leakage characteristics [10]. Here, the asymmetric leakage characteristic is dubbed *one-sided leakage*, which also occurs in other types of dynamic bit-cells.

A compounded 2T-GC bit-cell was proposed in [15] for soft error immunity. To achieve inherent per-bit error detection, the authors combined two GCs of the same type storing opposite logic values, supplemented with parity detection capabilities. Here we examine a different question of whether in all leakage circumstances, the readout alone (without further error detection) is sufficient to deduce the stored value.

3 The Compounded Bit-Cell: A Possible Remedy for the Limited DRT Problem

Let us consider an enhancement of the bit-cell in an attempt to obtain an infinite DRT, or at least extend it substantially. The bit-cell is shown in Fig. 4. It stores both Z and $\bar{\mathrm{Z}}$, where each polarity is stored in $\mathrm{N-GC}$ and P-GC. The N and P cells storing $\bar{\mathrm{Z}}$ are

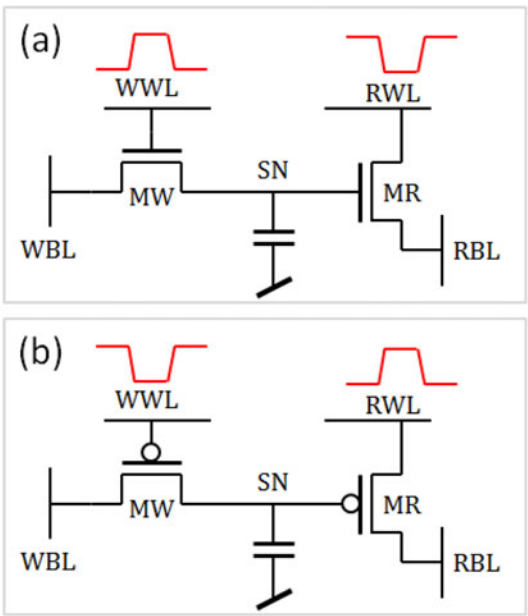

Fig. 2. 2T-GC, (a) $\mathrm{N-GC}$ and (b) $\mathrm{P-GC}$.

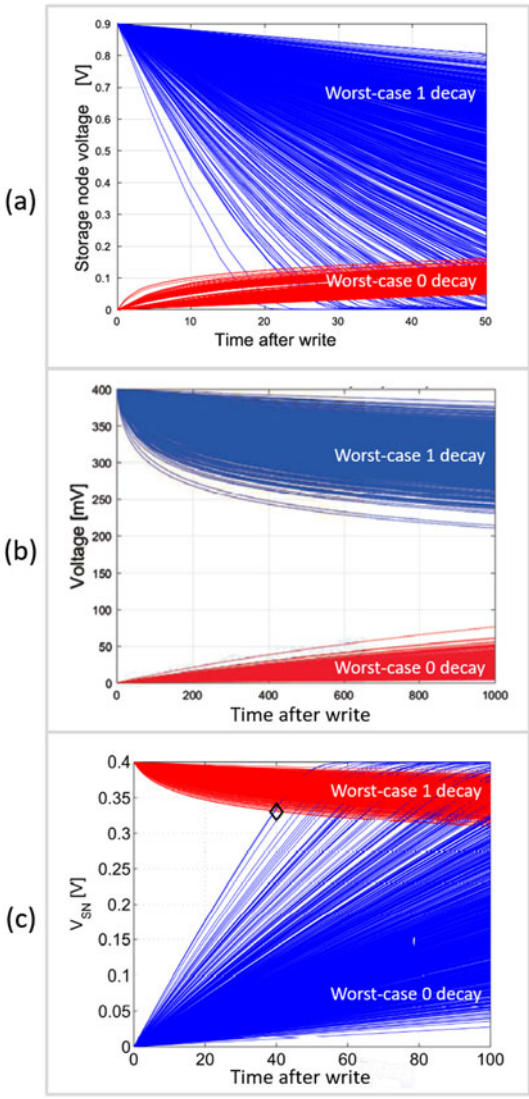

Fig. 3. Leakage characteristics of GCs, (a) 2T $\mathrm{N-GC}$ [9], (b) 5T $\mathrm{N-GC}$ [14] and (c) 2T $\mathrm{P-GC}$ [10].

denoted by $\bar{\mathrm{N}}$ and $\bar{\mathrm{P}}$, respectively. This gives rise to four storage nodes compounded in a single bit-cell. Although this may look expensive from an implementation standpoint, it would be worth it if the above DRT goal is met.

For the sake of the analysis and clarity, the one-sided leakage model is simplified by making the subsequent assumptions. The model is shown in Fig. 5, reflecting the leakage characteristics in Fig. 1.

1. During the time period after Z is written into the cell, WBL presents the worst-case scenario when its value is fixed for a sufficiently long time (either $\mathrm{WBL} = 0$ or $\mathrm{WBL} = 1$) oppositely to the value stored at the node, and
2. Data readout of the storage nodes is either the originally stored value Z if leakage did not occur or its complement $\bar{\mathrm{Z}}$ if the storage node leaked.

Let the bit-cell be read out after a sufficiently long time. It is unknown upfront which of the four internal storage nodes in Fig. 4 was subject to data leakage. We thus need to consider all the $16 = 2^4$ possibilities listed in (1), each of which represents a hypothesis that the nodes within the brackets have leaked their data.

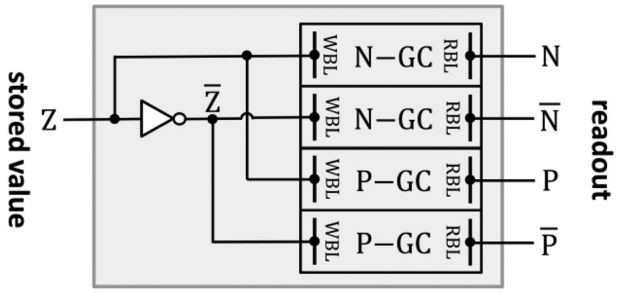

Fig. 4. Compounded bit-cell comprising four internal storage nodes.

$$\begin{aligned}&\{\emptyset\},\\&\{\mathrm{N}\},\{\bar{\mathrm{N}}\},\{\mathrm{P}\},\{\bar{\mathrm{P}}\},\\&\{\mathrm{N}\bar{\mathrm{N}}\},\{\mathrm{NP}\},\{\mathrm{N}\bar{\mathrm{P}}\},\{\bar{\mathrm{N}}\mathrm{P}\},\{\bar{\mathrm{N}}\bar{\mathrm{P}}\},\{\mathrm{P}\bar{\mathrm{P}}\},\\&\{\mathrm{N}\bar{\mathrm{N}}\mathrm{P}\},\{\mathrm{N}\bar{\mathrm{N}}\bar{\mathrm{P}}\},\{\mathrm{NP}\bar{\mathrm{P}}\},\{\bar{\mathrm{N}}\mathrm{P}\bar{\mathrm{P}}\},\\&\{\mathrm{N}\bar{\mathrm{N}}\mathrm{P}\bar{\mathrm{P}}\}\end{aligned}\tag{1}$$

This makes it possible to validate whether any of the above hypotheses conforms to the data observed by the readout. There are two cases:

1. The leakage hypothesis conforms to the readout, a case where the stored value Z can be deduced.
2. The leakage hypothesis and the readout are contradictory; hence, the leakage hypothesis is invalid for such readout.

To demonstrate this idea consider the following readout: $\mathrm{N}=1$, $\bar{\mathrm{N}}=0$, $\mathrm{P}=1$, and $\bar{\mathrm{P}}=1$, represented by the quadruple (1011), shown in line 11 of Table 1. Definitely the $\{\emptyset\}$ leakage hypothesis; namely that none of the nodes leaked, is invalid since it contradicts the readout $\mathrm{P}=\bar{\mathrm{P}}$, whereas initially $\mathrm{P}=\mathrm{Z}$ and $\bar{\mathrm{P}}=\bar{\mathrm{Z}}$, hence if leakage did not occur then there must be $\mathrm{P}\neq\bar{\mathrm{P}}$. The leakage hypothesis $\{\mathrm{N}\bar{\mathrm{N}}\mathrm{P}\bar{\mathrm{P}}\}$; namely that all nodes leaked, is also invalid. This follows since initially $\mathrm{P}\neq\bar{\mathrm{P}}$, and by the hypothesis both have leaked their value, so their readout still must be $\mathrm{P}\neq\bar{\mathrm{P}}$. This however contradicts the readout $\mathrm{P}=\bar{\mathrm{P}}$.

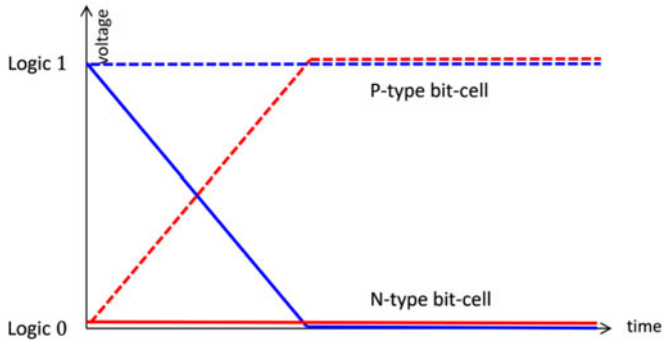

Fig. 5. Simplified leakage characteristics of a dynamic memory cell.

On the other hand, the $\{\bar{\mathrm{P}}\}$ leakage hypothesis is valid since it dictates that $\mathrm{P}=\bar{\mathrm{P}}$, which agrees with the readout. Additionally, the initial state of $\mathrm{N}\neq\bar{\mathrm{N}}$ is intact and agrees with $\mathrm{N}=\mathrm{P}$ (see Fig. 4). All in all, the $\{\bar{\mathrm{P}}\}$ leakage hypothesis conforms to the (1011) readout. By using similar logic, one can verify that all the remaining hypotheses in (1) are invalid. Since the $\{\bar{\mathrm{P}}\}$ leakage hypothesis is the one and only that conforms to the (1011) readout, the assertion that $\mathrm{Z}=1$ is conclusive as shown in line 11 of Table 1.

Whereas the above (1011) readout yielded the conclusive assertion $\mathrm{Z}=1$, let us examine another readout: $\mathrm{N}=1$, $\bar{\mathrm{N}}=1$, $\mathrm{P}=0$, and $\bar{\mathrm{P}}=1$, represented by the quadruple (1101) shown in line 13 of Table 1. Let us verify whether the assertion $\mathrm{Z}=1$ is valid. If it is valid, then the readout $\bar{\mathrm{N}}=1$ means that the initial $\bar{\mathrm{N}}=0$ has been $0\rightarrow1$ leaked. However, according to the leakage model in Fig. 5, an N-type node does not leak $0\rightarrow1$. So let us verify whether the assertion that $\mathrm{Z}=0$ is valid. If it is valid, then the readout $\mathrm{N}=1$ means that the initial $\mathrm{N}=0$ has been $0\rightarrow1$ leaked, which is also impossible for the same reason as before. To be more rigorous, one can validate all the leakage hypotheses in (1) to find that none conforms to the readout (1101), which is thus infeasible and cannot occur.

Since the bit-cell compounds four storage nodes, and each can be read out in 0 or 1, there are $16=2^4$ possibilities for the readout

TABLE 1
The Possible Readouts and Leakage Hypotheses of a Compounded Bit-Cell

	Readout				Leakage hypothesis															Conclusion				
	N	$\bar{N}$	P	$\bar{P}$	∅	N	$\bar{N}$	P	$\bar{P}$	N$\bar{N}$	NP	N$\bar{P}$	$\bar{N}$P	$\bar{N}\bar{P}$	P$\bar{P}$	N$\bar{N}$P	N$\bar{N}\bar{P}$	NP$\bar{P}$	$\bar{N}$P$\bar{P}$	N$\bar{N}$P$\bar{P}$	N	$\bar{N}$	P	$\bar{P}$	**Z**
0	0	0	0	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
1	0	0	0	1	X	X	●	X	X	X	X	X	X	X	X	X	X	X	X	X	0	1	0	1	**0**
2	0	0	1	0	X	●	X	X	X	X	X	X	X	X	X	X	X	X	X	X	1	0	1	0	**1**
3	0	0	1	1	X	X	X	X	X	X	X	●	●	X	X	X	X	X	X	X	undecidable				**1/0**
4	0	1	0	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
5	0	1	0	1	●	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	0	1	0	1	**0**
6	0	1	1	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
7	0	1	1	1	X	X	X	●	X	X	X	X	X	X	X	X	X	X	X	X	0	1	0	1	**0**
8	1	0	0	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
9	1	0	0	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
10	1	0	1	0	●	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	1	0	1	0	**1**
11	1	0	1	1	X	X	X	X	●	X	X	X	X	X	X	X	X	X	X	X	1	0	1	0	**1**
12	1	1	0	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
13	1	1	0	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
14	1	1	1	0	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				
15	1	1	1	1	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	cannot occur				

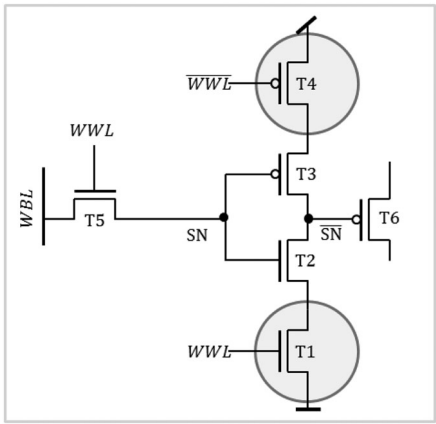

Fig. 6. Tristate inverter basic bit-cell.

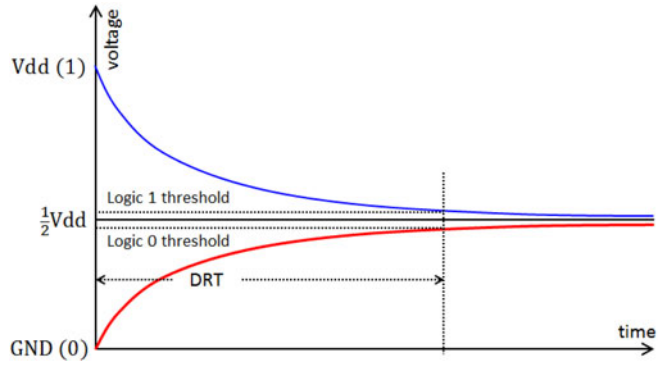

Fig. 7. Symmetric leakage in the tristate inverter bit-cell.

shown by Table 1. As already seen, some are feasible, yielding a conclusive assertion of Z, whereas others can never occur. There is still the case to consider where more than one of the hypotheses in (1) conforms to the readout.

So let us consider the readout (0011) shown in line 3 of Table 1, corresponding to $\mathrm{N} = 0$, $\bar{\mathrm{N}} = 0$, $\mathrm{P} = 1$, and $\bar{\mathrm{P}} = 1$. Let us consider the $\{\mathrm{N}\bar{\mathrm{P}}\}$ leakage hypothesis. If $\mathrm{Z} = 1$ was stored, by the leakage hypothesis and by the $1 \rightarrow 0$ N-type leakage, $\mathrm{N} = 0$ and $\bar{\mathrm{N}} = 0$ are valid readouts. Similarly, since the P-type leaks $0 \rightarrow 1$, $\mathrm{P} = 1$, and $\bar{\mathrm{P}} = 1$ are valid readouts. All in all, the $\mathrm{Z} = 1$ could be asserted, unless there is another leakage hypothesis yielding the $\mathrm{Z} = 0$ assertion. This is indeed what happens for the leakage hypothesis $\{\bar{\mathrm{N}}\mathrm{P}\}$ where $\bar{\mathrm{N}}$ leaked $1 \rightarrow 0$ and P leaked $0 \rightarrow 1$. Consequently, though the readout (0011) is feasible, there is no way to conclude what Z was stored as shown in Table 1 on line 3.

Table 1 summarizes all the conclusions that can be made on Z without knowing anything about what leakage occurred in the internal nodes of the compounded cell. An X means that this particular readout cannot occur for the specified leakage hypothesis. A red circle means a valid readout. The conclusion column summarizes whether the readout can occur at all, and if so, what can be concluded about Z. A readout which is all Xs cannot occur. A readout matching a unique leakage hypothesis yields a conclusive assertion of Z. The yellow row is the case where the readout could occur for two different leakage hypotheses yielding contradictory Z assertions.

4 A New Complementary Bit-Cell

Section 3 showed the existence of an inherent incompleteness regardless of which combination of nodes are used and polarities are stored. There is a readout case where after a sufficiently long time the value stored at the bit-cell cannot be asserted.

Inspired by the dual-polarity idea, below we propose a new promising CMOS compatible bit-cell circuit architecture [8]. Unlike the aforementioned compounds where the storage nodes are independent of each other, the new circuit captures both polarities in two mutually dependent storage nodes. Though refreshing may still be required, the circuit presents two crucial advantages. First, its DRT is considerably longer than existing eDRAM CMOS bit-cells. Moreover, due to its *two-sided* symmetric leakage, its DRT can theoretically be infinite. Second, the mutual dependence of the storage nodes makes it possible to refresh the entire memory array in a single cycle, thus almost totally nullifying the unavoidable performance loss occurred by row-by-row ordinary refreshing.

Returning to the 2T-GCs in Fig. 2, the main reason for its data leakage at SN is the poor isolation from WBL. Although 5T-GC [14] has improved DRT substantially, the direct connection of SN to WBL through a pass-gate still remains the main cause of leakage. To avoid the direct connection to the storage node and improve its isolation, we suggest using a tristate inverter as shown in Fig. 6. Although SN is still connected directly to WBL, $\overline{\mathrm{SN}}$ is isolated and a permanent WBL value opposite of SN should not affect $\overline{\mathrm{SN}}$. Moreover, the symmetric connection of $\overline{\mathrm{SN}}$ to Vdd (1) and GND (0) through the pullup branch T3 T4 and the pulldown branch T1 T2 ensures two-sided symmetric asymptotic leakage towards $\mathrm{Vdd}/2$. Unlike the typical one-sided $1 \rightarrow 0$ and $0 \rightarrow 1$ leakage, where after a sufficiently long time the bit level swaps, the symmetric leakage lets us distinguish between 1 and 0, as illustrated in Fig. 7.

The tristate leakage behavior is fundamentally different from the behavior of the typical dynamic bit-cells shown in Fig. 1. Though theoretically it could yield an infinite DRT, the finite margins of the sense amplifiers reading out the stored data imply a finite DRT as illustrated in Fig. 7.

It is important to note that although the circuit in Fig. 6 is comprised of six transistors, the bit-cell per se only has four transistors, where T6 is used for readout similar to MR in Fig. 1. The encircled T1 and T4 are shared by all the bits of a memory row as illustrated in Fig. 8. The two shared transistors should be sized appropriately to supply the current drawn at the time the underlying word is written. A sizing discussion is beyond the scope of this brief.

5 Self-Refreshing

The coupling of the storage nodes SN and $\overline{\mathrm{SN}}$ enables self-refreshing of the bit-cells as explained below. The refreshing of the dynamic memories (1T1C, GC) requires copying the bit data into a register, aka a refreshing buffer, and then writing it back into the

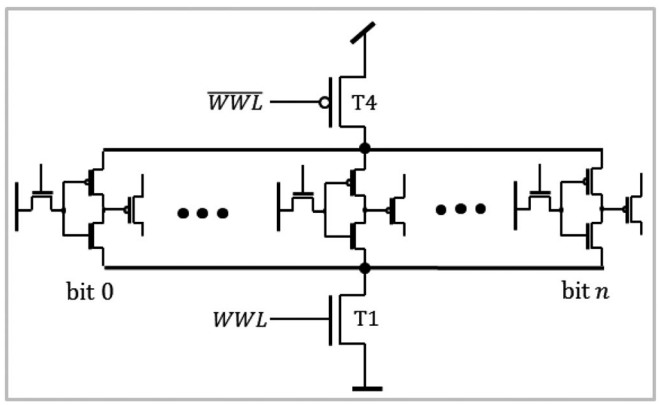

Fig. 8. A memory word sharing of its write transistors.

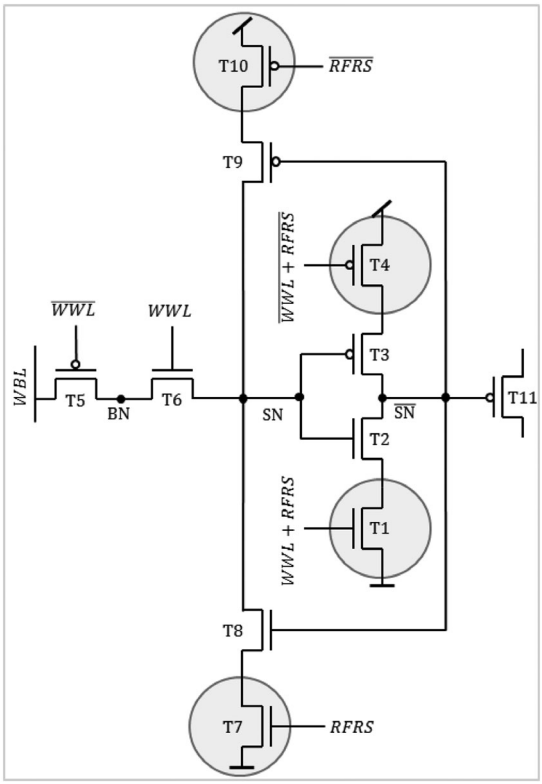

Fig. 9. Fully featured self-refreshable tristate inverter basic bit-cell.

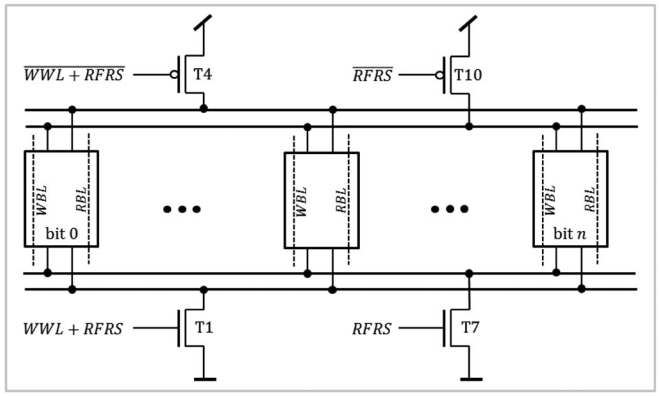

Fig. 10. A memory word sharing of its write and refreshing transistors.

bit. Ordinary refreshing has two major drawbacks: it loses clock cycles proportionally to the number of words in the refreshable unit (e.g., memory bank), and it also consumes high power. Self-refreshing avoids these performance and power overheads. Instead, it uses the complementary storage node $\overline{\mathrm{SN}}$ to self-refresh the counterpart node SN. To the best of our knowledge self-refreshing of a bit by its internals is novel and does not exist in any dynamic memory to date.

Note that the self-refreshing notion proposed here should not be confused with the self-refreshing notion related to ordinary DRAM operations, where the DRAM controller initiates a line-by-line refreshing of the underlying word-lines [16].

The fully featured self-refreshable bit-cell is illustrated in Fig. 9, where the encircled transistors are not part of the bit-cell per se but rather are shared by the entire underlying word in a similar fashion as in Fig. 8.

Before elaborating on the self-refreshing, let us first closely examine the write of the bit-cell. Although for writing WBL into SN a single pass-gate as T5 in Fig. 6 suffices, it turns out that the addition of a complementary transistor to the write path like T5 T6 in Fig. 9 slows down SN leakage drastically.

Let $\mathrm{WBL} = 1$. Setting $\mathrm{WWL} = 1$ connects the paths T5 T6. The writing cycle completes once $\mathrm{WWL} = 0$, disconnecting the paths T5 T6, a time when both $\mathrm{BN} = 1$ and $\mathrm{SN} = 1$. Assume the worst case where once the write cycle completes, $\mathrm{WBL} = 0$ permanently. The values of SN and BN decrease towards 0 due to the sub-threshold leakage through T5 and T6, whereas BN decreases faster than SN due to the voltage divider T5 T6. The node BN can be thought of as the source of T5 which is P-type and the voltage in BN is higher than WBL voltage. Since $\overline{\mathrm{WWL}} = 1$ T5 satisfies $V_{GS} > 0$, thus decreasing the sub-threshold leakage through T5. As the decrease of BN voltage is likely to continue, the increase of T5 V_{GS} further suffocates its leakage.

For the opposite case where both $\mathrm{BN} = 0$ and $\mathrm{SN} = 0$ after write, let us consider the worst case where once the write cycle completes $\mathrm{WBL} = 1$ permanently. The values of SN and BN increase towards 1, whereas BN increases faster than SN due to the voltage divider T5 T6. Node SN can be thought of as the source of T6 which is N-type and the voltage in BN is higher than SN voltage. Since $\mathrm{WWL} = 0$ T6 satisfies $V_{GS} < 0$, thus decreasing the sub-threshold leakage through T6. As the increase of SN voltage is likely to continue, the decrease of T6 V_{GS} further suffocates its leakage.

Next we describe how self-refreshing operates. Let $\mathrm{SN} = 1$ and $\overline{\mathrm{SN}} = 0$, and assume the worst case where $\mathrm{WBL} = 0$ permanently, causing steady $1 \rightarrow 0$ leakage of SN. As shown in Fig. 7, the DRT period elapses after the latest write of the bit-cell calls for refreshing. This is handled by setting $\mathrm{RFRS} = 1$ in Fig. 9. A connection of SN to V_{dd} through the T9 T10 path is established and the recharge of SN to 1 takes place. At the same time RFRS connects the T1 T2 pulldown path of the tristate inverter to GND, discharging $\overline{\mathrm{SN}}$ to 0, thus completing the bit refreshing.

Consider the opposite case where $\mathrm{SN} = 0$ and $\overline{\mathrm{SN}} = 1$, and assume the worst case where $\mathrm{WBL} = 1$ permanently, causing steady $0 \rightarrow 1$ leakage of SN. By enabling the signal RFRS, connection of SN to GND through the T7 T8 path is established and the discharge of SN to 0 takes place. At the same time RFRS connects the T3 T4 pullup path of the tristate inverter to V_{dd}, recharging $\overline{\mathrm{SN}}$ to 1, thus completing the bit refreshing.

A whole memory word is shown in Fig. 10, where the encircled transistors of Fig. 9 are shared by all the bits, sized appropriately to supply the current drawn by refreshing and writing. The sizing details are beyond the scope of this brief.

Ordinary refreshing must take place row-by-row, blocking the access between the CPU and memory within a DRT period for the number of clock cycles proportional to the memory size. It was shown in [7] that a performance degradation of 50% can occur, depending on the read/write instructions ratio and their frequency. This is completely avoided by self-refreshing.

Self-refreshing nullifies performance loss by simply enabling the global signal $RFRS$ for the entire rows of the memory unit. Such one-shot refreshing blocks the CPU access for just a single cycle. It may obviously draw a very high supply current, causing a considerable V_{dd} drop. In addition to appropriate sizing of the shared transistors it is also possible to refresh smaller memory chunks in a few cycles. This would still take far less time than that required by row-by-row refreshing.

Another advantage of the self-refreshing feature is the possibility to considerably speed up the bit-cell readout. The leakage of the data stored at the bit-cell usually causes considerable slowdown of data readout. A readout boost is obtained by simply enabling the signal $RFRS$ for the underlying memory word together with its read enable signal. Applying self-refreshing simultaneously accelerates the readout significantly.

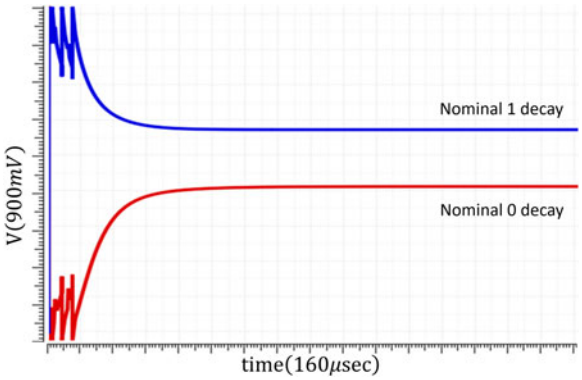

Fig. 11. 28nm bit-cell $\overline{\mathrm{SN}}$ voltage obtained by SPICE simulation.

6 Experimental Results

Fig. 11 shows the $\overline{\mathrm{SN}}$ voltage change over time obtained by a SPICE simulation at a nominal typical-typical, $\mathrm{Vdd} = 0.9\mathrm{V}, T = 85^{\circ}\ \mathrm{C}$, corner in 28nm process technology. It is shown that though the bit-cell values leaked, they stayed above and below $\mathrm{Vdd}/2$ in accordance with Fig. 7. The spikes on the left side are simply the few instances of refreshing that took place before the values of the cell were let to leak for a long time period.

Fig. 12 simulates writing into the bit-cell an opposite value than what it stored. The blue curve shows the bit-cell when its previous value was 1 and then the 0 value was written. Its value indeed swapped. Similarly, the red curve shows the bit-cell when its previous value was 0 and then 1 value was written.

Fig. 13 shows the bit-cell in Monte Carlo simulations of various corners and device mismatches to verify its behavior under the worst operating conditions. The apparent infinite DRT shown in Figs. 11 and 12 for a nominal corner can no longer be expected in worst-case. In fact, after $50\mu\mathrm{sec}$ the worst-case gap becomes narrow, which depending on the characteristics of the sense amplifier used for the readout, may dictate the actual DRT.

Note that when comparing the DRT to GC we used the same standard 28nm technology as the authors in [9]. Figs. 11 and 13 show that the 1 and 0 logic levels do not cross each other, whereas the GCs in Fig. 3 do. However, it is important to note that the authors of [17] achieved a considerably longer DRT of 1.6msec using a 4T-GC implemented in FD-SOI 28nm technology.

Self-refreshing comes with no energy overhead. Averaging power along the DRT period, dynamic energy consumption comprises two evenly contributing factors: the capacitance (area) of the storage node SN and the refreshing register. Whereas our DRT and bit-cell SN area is in the ballpark of [17] (gates and diffusion nodes count), self-refreshing avoids the latter factor.

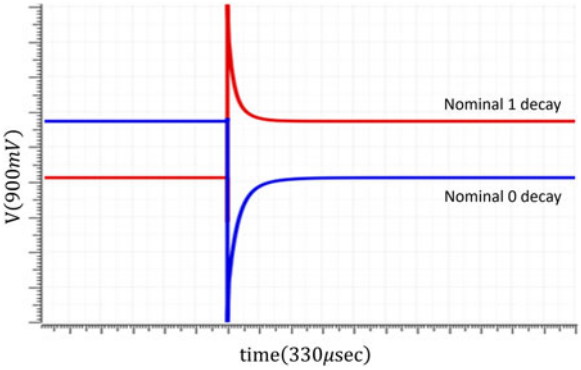

Fig. 12: Simulation of swapping bit-cell values.

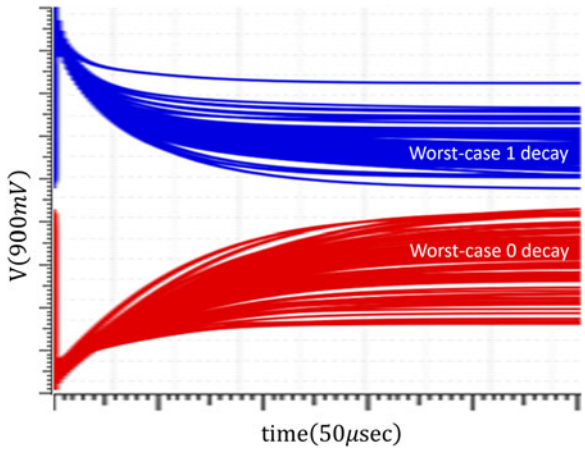

Fig. 13. Bit-cell in Monte Carlo simulations of various corners.

7 Conclusion

This brief proved that compounding various types of storage nodes of asymmetric leakage in a single CMOS eDRAM bit-cell, storing the datum and its complement, cannot ensure proper readout in an unrestricted time period after that datum was written. In other words, limited DRT is unavoidable.

By adopting the idea of dual-polarity complementary storage nodes we presented a new CMOS compatible eDRAM self-refreshable bit-cell circuit architecture based on a tristate inverter. The proposed bit-cell yields considerably extended DRT, which is practically finite but theoretically infinite. The dual-polarity property enabled the refreshing of an entire memory array in a single clock cycle, thus almost totally nullifying the unavoidable performance loss occurred by row-by-row ordinary refreshing. This brief focused on the circuit architecture and the implications for the entire memory. The detailed circuit design is left for future work.

Acknowledgments

The authors wish to thank Mr. Rafi Hana for his helpful comments and advices. The authors are also grateful for the useful comments by the anonymous reviewers which helped to improve this paper. The second author conducted this research while being with Prof. Atsushi Takahashi Lab at the Tokyo Institute of Technology.

References

[1] H. B. Calhoun and A. P. Chandrakasan, "A 256-kb 65-nm sub-threshold SRAM design for ultra-low-voltage operation," *IEEE J. Solid-State Circuits*, vol. 42, no. 3, pp. 680–688, Mar. 2007.

[2] C. K. Chun, P. Jain, T.-H. Kim, and C.. H. Kim, "A 667 MHz logic-compatible embedded DRAM featuring an asymmetric 2T gain cell for high speed on-die caches," *IEEE J. Solid-State Circuits*, vol. 47, no. 2, pp. 547–559, Feb. 2012.

[3] P. Meinerzhagen, A. Teman, R. Giterman, N. Edri, A. Burg, and A. Fish, *Gain-cell Embedded DRAMs For Low-Power VLSI Systems-on-Chip*. Berlin, Germany: Springer, 2018.

[4] J. Liu, B. Jaiyen, Y. Kim, C. Wilkerson, and O. Mutlu, "An experimental study of data retention behavior in modern DRAM devices: Implications for retention time profiling mechanisms," *ACM SIGARCH Comput. Archit. News*, vol. 41, no. 3, pp. 60–71, 2013.

[5] B. Jacob, D. Wang and N. Spencer, *Memory Systems: Cache, DRAM, Disk*. San Mateo, CA, USA: Morgan Kaufmann, 2010.

[6] A. Kazimirsky and S. Wimer, "Opportunistic refreshing algorithm for eDRAM memories," *IEEE Trans. Circuits Syst. I, Regular Papers*, vol. 63, no. 11, pp. 1921–1932, Nov. 2016.

[7] B. Frankel, R. Herman and S. Wimer, "Queuing-based eDRAM refreshing for ultra-low power processors," *IEEE Trans. Comput.*, vol. 67, no. 9, pp. 1331–1340, Sep. 2018.

[8] B. Frankel and S. Wimer, "A dynamic memory bit cell based on tristate inverter," US Patent 63/138,820, Jan. 19, 2021.

[9] R. Giterman, Y. Weizman and A. Teman, "Gain-cell embedded DRAM-based physical unclonable function," *IEEE Trans. Circuits Syst. I, Regular Papers,* vol. 65, no. 12, pp. 4208–4218, Dec. 2018.
[10] N. Edri, P. Meinerzhagen, A. Teman, A. Burg, and A. Fish, "Silicon-proven, per-cell retention time distribution model for gain-cell based eDRAMs," *IEEE Trans. Circuits Syst. I, Regular Papers,* vol. 63, no. 2, pp. 222–232, Feb. 2016.
[11] P. Meinerzhagen, A. Teman, R. Giterman, A. Burg, and A. Fish, "Exploration of sub-VT and near-VT 2T gain-cell memories for ultra-low power applications under technology scaling," *J. Low Power Electron. Appl.,* vol. 3, no. 2, pp. 54–72, 2013.
[12] K. Roy, S. Mukhopadhyay, and H. Mahmoodi-Meimand, "Leakage current mechanisms and leakage reduction techniques in deep-submicrometer CMOS circuits," *Proc. IEEE,* vol. 91, no. 2, pp. 305–327, Feb. 2003.
[13] A. Sanyal, A. Rastogi, W. Chen and S. Kundu, "An efficient technique for leakage current estimation in nanoscaled CMOS circuits incorporating self-loading effects," *IEEE Trans. Comput.,* vol. 59, no. 7, pp. 922–932, Jul. 2010.
[14] R. Giterman, A. Teman, and A. Fish, "A 11.5 pW/bit 400mV 5T gain-cell eDRAM for ULP applications in 28nm FD-SOI," in *Proc. IEEE SOI-3D-Subthreshold Microelectronics Technol. Unified Conf.,* 2017, pp. 1–3.
[15] R. Giterman, L. Atias, and A. Teman, "Area and energy-efficient complementary dual-modular redundancy dynamic memory for space applications," *IEEE Trans. Very Large Scale Integration Syst.,* vol. 25, no. 2, pp. 502–509, Feb. 2017.
[16] B. Oh, N. Abeyratne, J. Ahn, R. G. Dreslinski, and T. Mudge, "Enhancing dram self-refresh for idle power reduction," in *Proc. Int. Symp. Low Power Electron. Des.*, 2016, pp. 254–259.
[17] R. Giterman, A. Fish, A. Burg and A. Teman, "A 4-transistor nMOS-only logic-compatible gain-cell embedded DRAM with over 1.6-ms retention time at 700 mV in 28-nm FD-SOI," *IEEE Trans. Circuits Syst. I, Regular Papers,* vol. 65, no. 4, pp. 1245–1256, Apr. 2018.