

A sensing circuit for single-ended read-ports of SRAM cells with bit-line power reduction and access-time enhancement

T. Heselhaus and T. G. Noll

Chair of Electrical Engineering and Computer Systems, RWTH Aachen University, 52062 Aachen, Germany

Abstract. The conventional sensing scheme of single-ended read-only-ports as integrated in 8T-SRAM cells suffers from low performance compared to double-ended complementary sensing schemes. In the proposed sensing scheme the pre-charge voltage of the single-ended read-bit-line is set to a level above the threshold voltage of the sensing device with an adjustable margin. This margin is minimized to speed up the read access on the one hand and kept large enough to provide a sufficient bit-line noise margin on the other hand. The pre-charge voltage level of the proposed sensing circuit tracks the threshold voltage of the sensing device under process variations in order to maintain a minimum required bit-line noise margin. To avoid unnecessary bit-line discharging, the proposed sensing scheme employs a modified 8T-SRAM cell. Compared to the conventional 8T-SRAM cell, the read port of the proposed cell provides a virtual ground line running in parallel to the bit-lines. An internal driver of the sensing circuit releases the virtual ground line during the evaluation period to prevent the charge dissipation resulting in a raised voltage level. The reduced pre-charge level and the increased virtual ground lead to a reduced bit-line voltage swing and thus a bit-line power reduction. Access time, energy dissipation, and noise margin of the proposed sensing circuit are compared with conventional sensing circuits from the literature for different numbers of memory cells connected to the bit-line. It is shown, that for a specific number of memory cells per bit-line the proposed circuit achieves fastest access time at low power operation.

1 Introduction

Single-ended read sensing schemes will become more important in memory architectures of SRAM cells with additional read-only-ports to overcome the reliability problems related to standard 6T-SRAM cells. Though such single-ended ports require only one bit-line, the main power dissipation originates from the full-swing bit-line sensing schemes.

The intention of this work is to set the initial bit-line voltage close to the threshold level of the individual local sensing device and stop the discharging, once the local sensing has detected the data bit, such that no more than necessary of the bit-line charge is dissipated. The expected side effect is a performance upgrade, since the initial bit-line voltage resides close to the sensing threshold level, such that this level can be reached in a shorter evaluation time.

Section 2 presents a modified memory cell, which is required to enable the bit-line swing reduction. In Sect. 3 the proposed sensing circuit is described. In Sect. 4 the proposed sensing scheme is verified with simulations and compared with other sensing circuits from the literature.

2 The proposed 8T-SRAM cell

Figure 1 shows the conventional (Chang et al., 2005) and the proposed 8T-SRAM cell. Both cell variants consist of a conventional cross-coupled inverter-pair connected via two access transistors to a complementary bit-line pair p and \bar{p} , which represents the write-port for storing data into the cell by activating the storage-word-line s . The read-only-port consists of two series connected transistors for single-ended read-out with a separate read-word-line r .

Area and most of the wiring of the proposed cell is identical to the conventional cell, however the proposed cell connects the source of the read-only-port to an additional virtual ground line v instead of V_{SS} . This additional virtual ground



Correspondence to: T. Heselhaus
(heselhaus@eecs.rwth-aachen.de)

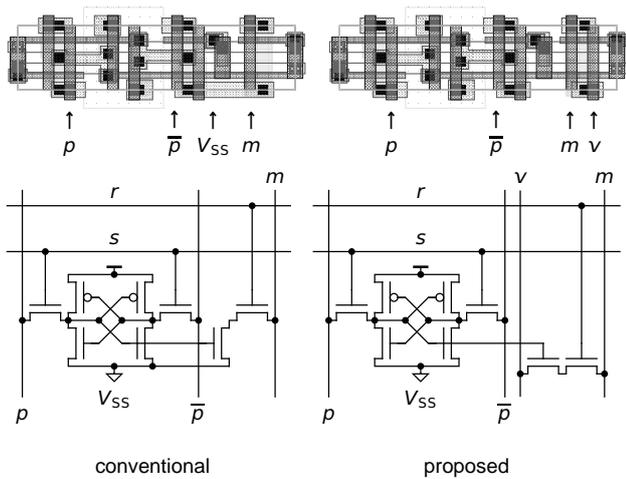


Fig. 1. Conventional and proposed 8T-SRAM cell as schematic and layout view. The additional virtual ground line is labeled as v .

line runs in parallel to the read-bit-line m and can be easily placed in the thin layout cell without any area penalty. The intention of this additional virtual ground line is to reduce the bit-line swing and power dissipation. The functionality of this virtual ground line will be described in the next section.

3 Pre-charge and sensing circuit

Figure 2 shows the pre-charge and sensing circuit of a single-ended bit-line for one column of memory cells. Only the read-only-port of one memory cell is shown, however, there are N memory cells connected to the bit-lines m and v . The basic idea of the pre-charge circuit is to pre-charge the bit-line m to a defined voltage level above the threshold-level of the sensing transistor M3. Because the drain current of M3 is used to pre-charge the bit-line, M3 should itself turn off as the pre-charge voltage is reached. The pre-charge process is described as follows.

During the pre-charge-phase ($\bar{\phi} = 0$) the word-line is deactivated ($r = 0$) and the sensing is disabled ($\bar{e} = 1$). For an initially discharged bit-line m the gate potentials of M3 and M4 are both zero and the bit-line is charged by the series connected transistors M3 and M4. As the gates of M5 and M6 are at zero, these transistors establish a voltage divider of the actual bit-line voltage. The gate of M3 is connected to the divided voltage u , which is a fraction of the bit-line voltage m . While the bit-line voltage rises, the tapped gate potential u rises as well. The pre-charge process of the bit-line stops, if u rises above $V_{DD} - |V_{th}|$, such that M3 enters the off-region. Hence the final pre-charge voltage $V_{m,max}$ of the bit-line has risen in excess of the sensing threshold level $V_{DD} - |V_{th}|$ of M3. This overshoot can be adjusted by choosing the appropriate dimensioning of M5 and M6.

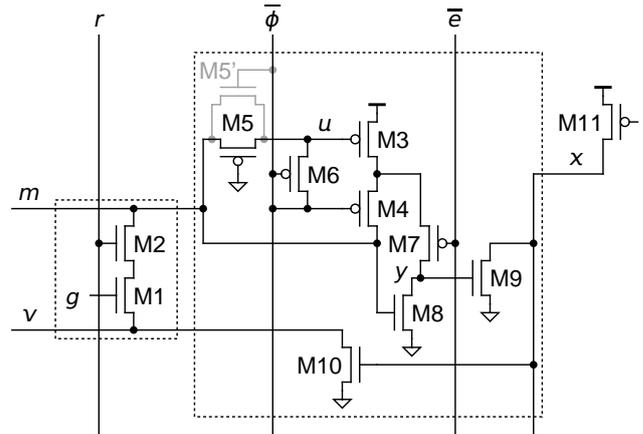


Fig. 2. Schematic of the proposed single-ended read sensing and pre-charge circuit for one memory column and domino connected digit line multiplex.

The intention of the overshoot is to achieve the required noise margin on the bit-line, which must be maintained even under process variations. With the feed-back loop for pre-charging the bit-line via the transistors M3, M4, M5 and M6, the voltage overshoot of $V_{m,max}$ tracks the sensing threshold level of M3. This tracking effect is verified by simulations in Sect. 4.

In fact, in the proposed sensing scheme the signal is detected with two successive sensing stages, realized by M3 and M9. It behaves similar to an inverter sense circuit followed by a domino NMOS pull-down device. However, the nodes x , y and v in the sensing circuit have to be pre-charged or clamped before entering the evaluation phase. During the pre-charge-phase the bit-line m rises to $V_{m,max}$ and shortly M8 begins to discharge the internal node y such that the main output stage M9 turns off. The drain of M9 is connected to a common node x , which could be a global bit-line in a subdivided bit-line memory architecture or a common digit line in some other hierarchical memory architecture (e.g. a column or data word multiplex output). Here, this output x is assumed to be domino-connected, thus x requires an additional pre-charge device M11, which is not part of the sensing circuit. However, M11 pre-charges x to V_{DD} such that M10 turns on and clamps the virtual ground line v to V_{SS} . The circuit is now prepared for sensing.

The pre-charge phase is completed by reverting $\bar{\phi}$ to V_{DD} which turns off M4. Since M6 is connected as a diode, M6 turns off as well, the voltage divider built by M5 and M6 is now no longer active and u follows the bit-line voltage via M5. The transistor M5' may be added to afford a close coupling between m and u over a wide range of bit-line voltages. Actually both M5' and M5 together build up a transmission gate to connect m to the gate of M3. In some cases where M5 is strong enough for a close coupling between m and u , M5' might be omitted. However, in this work M5' is included

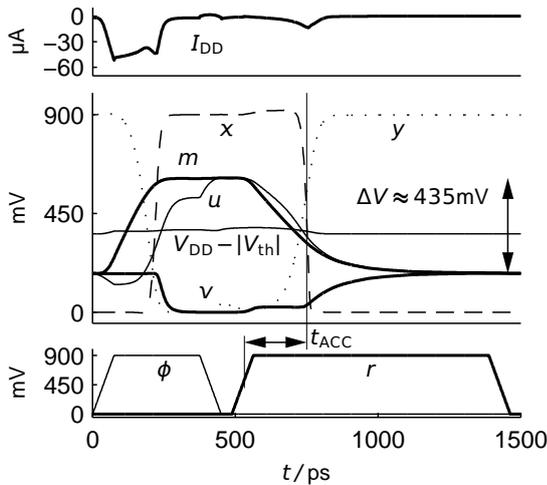


Fig. 3. Simulated waveform for a read-cycle of the proposed sensing circuit (slow corner, 125°C, $V_{DD} = 900\text{mV}$).

in each power and performance simulation and comparison. The transistor M5' is minimally dimensioned.

Additionally, \bar{e} is now tied to zero, and M7 connects the sensing transistor M3 to the internal node y . Any leakage of M3 due to the near threshold-level pre-charged bit-line m and node u is compensated now by the weakly dimensioned transistor M8, such that y is kept to V_{SS} . The evaluation-phase starts by activating the word-line r . Depending on the stored data g , the bit-line remains pre-charged or is discharged if $g = V_{DD}$ via M2 and M1 of the memory cell and M10 of the sensing circuit. The gate potential u of M3 falls and M3 now directly turns on, which leads to a steep rising edge on the internal node y . The rising edge on y activates M9, which now discharges the common data output node x .

The power reduction feature is induced by the reduced pre-charge level and the discharge inhibition with the additional virtual ground line v of the proposed 8T-SRAM cell and transistor M10 of the sensing circuit. Since M10 is connected to the data output x , which gets zero if the selected memory cell discharges the bit-line m , M10 turns off and inhibits further charge dissipation to V_{SS} as soon the data bit is detected. Because the bit-lines m and v are designed similar in the layout view, the parasitic capacitances of both bit-lines are similar. Thus, once M10 turns off, the remaining charge on the bit-line m will be balanced between m and v . The final bit-line potential of m would therefore settle to nearly half of the bit-line potential, when the data bit was detected. The final potential depends on the ratio of the parasitic capacitances of m and v .

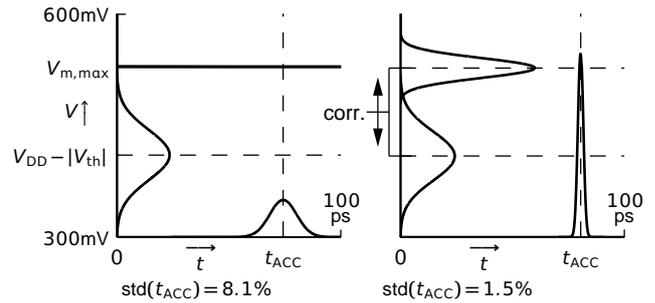


Fig. 4. Verification of the sensing threshold level tracking. The left diagram shows the resulting access time profile with a fixed bit-line pre-charge voltage, while the right diagram shows the results when the sensing circuit includes the voltage divider M5/M6 for the tracking (Monte Carlo simulations, 25°C, $V_{DD} = 900\text{mV}$).

4 Simulation results

The sensing circuit was simulated for a 40nm general purpose bulk CMOS technology with a nominal supply voltage of 900mV. All circuit simulations are based on transistor schematics. Extracted parasitics from a layout view are added to the bit-lines and to the virtual ground line corresponding to the connected number of cells $N \in \{8, 16, 32, 64, 128\}$. Internal nodes of the sensing circuit are loaded with estimated parasitic capacitances rated between 0.4fF and 0.7fF depending on the assumed wiring complexity. The transistor dimensioning of the proposed sensing circuit was basically designed with the minimal allowed gate width of $w = 120\text{nm}$. Only the driving/sensing transistors M3, M4, M9, and M10 have a gate width of $w = 480\text{nm}$.

Simulation using slow corner device models at 125°C and a minimal supply voltage of 900mV are used to verify the function of the proposed sensing circuit under worst case timing conditions. Figure 3 shows the simulated pre-charge and evaluation-phase of a read cycle with a stored “1” on the selected memory cell ($g = V_{DD}$). The achieved bit-line swing ΔV is approximately 435mV, a little bit less than $V_{DD}/2$ such that bit-line energy savings of 50% compared to a full swing bit-line sensing scheme can be expected.

Tracking the pre-charge voltage according to the sensing threshold level of M3 under process variations was functionally verified by Monte Carlo simulations. For this case process variations apply to M3 only, while all other transistors operate in the typical corner. Figure 4 shows the probability density functions (PDF) of sensing threshold level variations and the resulting access time profile without tracking on the left side and with tracking on the right side. On the ordinate the mean sensing threshold level V_{th} , the mean pre-charge level $V_{m,max}$, and their PDFs are plotted. The abscissas show the mean access times t_{ACC} and their PDFs for both cases. The standard deviation of the access time PDF was reduced from 8.1% to 1.5%.

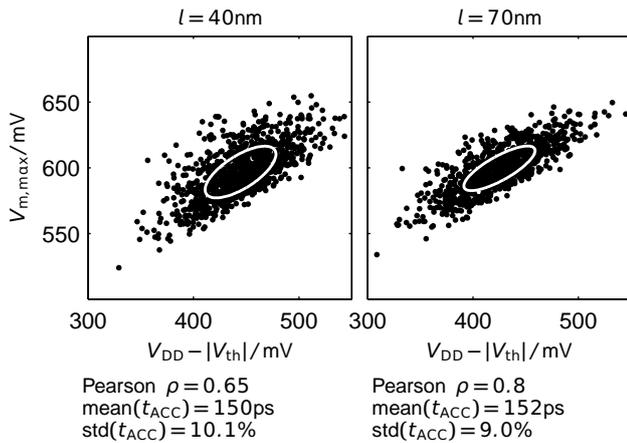


Fig. 5. Simulated pre-charge level correlation to M3-threshold-level variations of 1000 Monte Carlo runs for two different gate lengths of the voltage divider M5/M6 (25°C, $V_{DD} = 900\text{mV}$).

When considering process variations of all transistors of the sensing circuit this also affects the fraction of the voltage divider M5/M6. Figure 5 shows the simulated correlation of the sensing threshold level variation to the pre-charge level variation. The left side of Fig. 5 shows the voltage level distributions, where all transistors of the sensing circuit in Fig. 2 are designed with a gate lengths of $l_{M5} = l_{M6} = 40\text{nm}$. Since device mismatches of the voltage divider transistors M5 and M6 disturb the correlation of the sensing threshold level to the pre-charge level, a reduction of the variability of M5 and M6 can improve the correlation. Choosing a gate length of $l_{M5} = l_{M6} = 70\text{nm}$ for the voltage divider M5/M6 improves the correlation coefficient to $\rho = 0.8$ compared to $\rho = 0.65$ (for $l_{M5} = l_{M6} = 40\text{nm}$). Reducing the variability of the voltage divider reduces the standard deviation of the access time from 10.1% to 9%, while the mean access time remains nearly constant.

Increasing the bit-line noise margin was simulated by enlarging the gate length of M5 up to $l_{M5} = 420\text{nm}$ for a series of Monte Carlo simulations, while keeping the gate length of M6 fixed to $l_{M6} = 70\text{nm}$. The results shown in Fig. 6 exhibit a similar distribution of the pre-charge level in correlation to the sensing threshold voltage like depicted on the right diagram in Fig. 5. However, the dot clouds are shifted to higher pre-charge levels with increasing gate length l_{M5} . The ellipses inside the dot clouds denote the standard error ellipse of the two dimensional density function. The resulting improvement of the mean bit-line noise margin with increasing the gate length $l_{M5} = 70\text{nm} \rightarrow \{140\text{nm}, 280\text{nm}, 420\text{nm}\}$ are respectively $\Delta V_{BL, \text{margin}} = \{35\text{mV}, 100\text{mV}, 150\text{mV}\}$.

The proposed sensing circuit was compared to the following sensing circuits (refer to Table 1). As explained for the proposed sensing circuit, the sensing circuits for comparison are simulated using transistor schematics with estimated parasitics for their internal nodes (0.4...0.7fF). The transis-

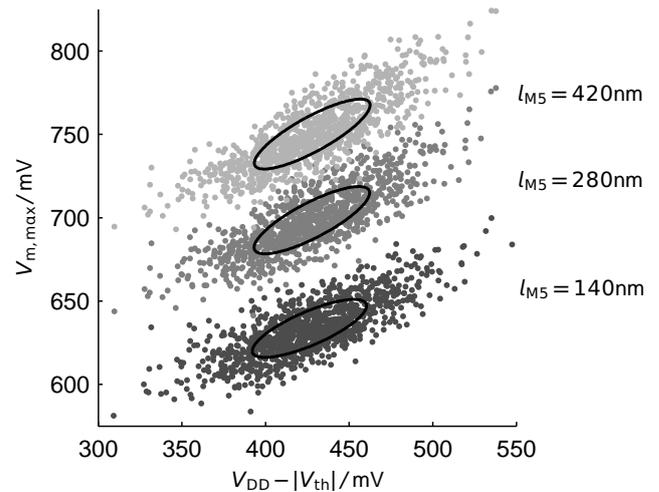


Fig. 6. Monte Carlo simulated pre-charge level for different l_{M5} in correlation to the M3-threshold-level ($l_{M6} = 70\text{nm}$, 1000 iterations, 25°C, $V_{DD} = 900\text{mV}$).

Table 1. Proposed and reference sensing circuits for comparison.

Designation	
A	The proposed sense circuit as shown in Fig. 2
B	A domino style sensing of the local bit-line m with a PMOS transistor, which implicitly implements the digit line multiplex (see Fig. 7) (Takeda et al., 2004)
C	Sensing with an inverter followed by an NMOS transistor for digit line multiplex (Cosemans et al., 2007). A similar sensing circuit implies a static CMOS NAND gate instead of an inverter for early digit line multiplex (Chang et al., 2007).
D	An AC coupled sense amplifier as PMOS replacement in a domino read path (Qazi et al., 2010). For the reduced bit-line swing due to word-line pulsing 400mV is assumed. To compare this circuit, the driving part for the global bit-line is included.
E	An AC coupled sense amplifier (Verma and Chandrakasan, 2009). For the reduced bit-line swing due to word-line pulsing 200mV is assumed.

tors are minimally dimensioned ($l = 40\text{nm}$, $w = 120\text{nm}$), except for pre-charge and other driving devices ($w = 480\text{nm}$). For any number of cells $N \in \{8, 16, 32, 64, 128\}$ the extracted parasitics from a column of conventional 8T-SRAM cells as shown in Fig. 1 are attached to the bit-line m of each reference circuit schematic. The circuit simulations of the reference circuits are applied for the same simulation corners, temperature and supply voltage as for the proposed sensing circuit. The cycle period was 2500ps. Figure 7 shows the reference circuits B and C as implemented for comparison. For the reference circuits domino-read and standard inverter the gate widths of the sensing transistor M3 and the

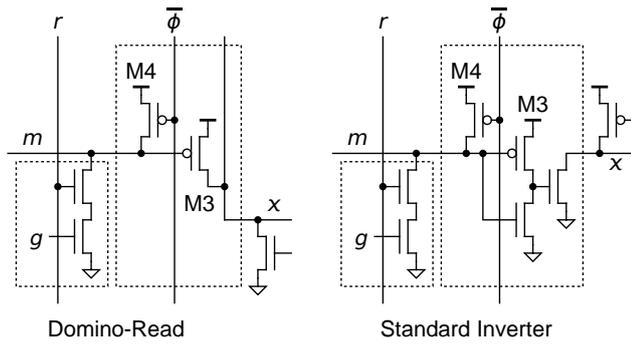


Fig. 7. Schematics of the domino-read (B) and inverter-read (C) sensing circuits for comparison.

Table 2. Simulated results for $N = 128$ cells (slow corner, 125°C , $V_{\text{DD}} = 900\text{mV}$).

	A	B	C	D	E
$t_{\text{ACC}} / \text{ps}$	300	408	512	401	325
$t_{\text{ACC,SA}} / \text{ps}$	82	45	53	94	126
$E_{\text{Sense}} / \text{fJ}$	5.08	1.12	2.21	5.58	9.95
$E_{\text{BL}} / \text{fJ}$	5.85	13.40	13.42	6.19	2.94
$E_{\text{tot}} / \text{fJ}$	10.93	14.53	15.62	11.77	12.89
$V_{\text{BL,margin}} / \text{mV}$	120	181	379	96	53
Transistor count ^a	10	3	5	13	14

^a without transistors of memory cells.

pre-charge transistor M4 were set to $w = 480\text{nm}$, all others were designed with minimal gate width of $w = 120\text{nm}$. Table 2 shows the simulated results of the proposed circuit A versus the reference circuits B...E for $N = 128$ cells. The access time t_{ACC} of the proposed circuit A is faster than B, C, D and E for $N = 128$ cells. The access time t_{ACC} consists of a bit-line dependent contribution $t_{\text{ACC,BL}}$ and an access time offset $t_{\text{ACC,SA}}$, which depends on the internal complexity of the sensing circuit. Apart from circuit B and C, the proposed circuit shows a smaller intrinsic access time $t_{\text{ACC,SA}}$ compared to circuit D and E.

The sense error immunity to bit-line noise was simulated with current pulses to inject charge to the bit-line during an evaluation period. Read evaluation faults are detected within an evaluation period of 2500ps . The tolerable bit-line noise $V_{\text{BL,margin}}$ is shown in Table 2 for each circuit.

Table 2 shows that the bit-line energy per cycle of circuit A is similar to circuit D. This results from nearly the same bit-line swing achieved in circuit A and D. The energy dissipation E_{Sense} of circuit A is less than that of circuit D and E. Since circuit B and C operate with full swing bit-lines, but dissipate the least energy E_{Sense} , there must be a crossover for a certain number of cells. Figure 8a shows that this crossover point can be identified at $N \geq 64$ cells.

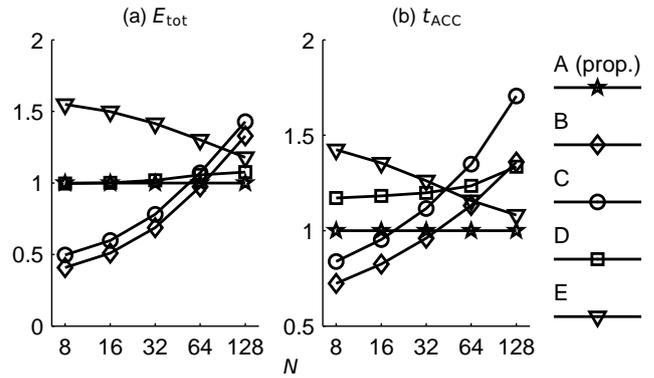


Fig. 8. Simulated t_{ACC} and E_{tot} for different number of connected cells N (normalized to the proposed circuit A, slow corner, 125°C , $V_{\text{DD}} = 900\text{mV}$).

Table 3. Comparison of the average access time and relative standard deviation of the proposed sensing circuit (with $l_{\text{M5}} = l_{\text{M6}} = 70\text{nm}$) for $N = 64$ using 1000 Monte Carlo simulations.

	A	B	C	D	E
mean (t_{ACC}) ^a	152 ps	167 ps	205 ps	169 ps	157 ps
std (t_{ACC}) ^a	9.0%	11.1%	9.5%	10.1%	10.6%
mean (t_{ACC}) ^b	153 ps	164 ps	204 ps	170 ps	158 ps
std (t_{ACC}) ^b	9.5%	14.7%	10.6%	10.6%	11.3%

^a constant temperature 25°C and supply voltage $V_{\text{DD}} = 900\text{mV}$.

^b normally distributed temperature and supply voltage variations, temperature $\in [-55^{\circ}\text{C} \dots 125^{\circ}\text{C}]$, $V_{\text{DD}} \in [810\text{mV} \dots 990\text{mV}]$.

A similar crossover point for the access time is located at $N \approx 40$ cells (see Fig. 8b). Simulations indicate smaller access times of circuit E for $N > 256$ cells. However, the proposed sensing circuit dissipates the least amount energy and shows the fastest access time compared to all reference circuits for $N \in [64 \dots 128]$.

Table 3 shows the mean access time over 1000 Monte Carlo iterations for each circuit and its standard deviation relative to the mean access time. The proposed sensing circuit provides the fastest mean access time and lowest access time variation compared to the reference circuits.

To implement the proposed sensing circuit in a hierarchical memory architecture, the sensing circuit on all but the last hierarchy level needs a small modification to provide the signaling scheme consisting of a digit line and a virtual ground line. The modification affects the output stage of the proposed sensing circuit and requires an additional internal inverter for driving the virtual ground switch. In Fig. 9 such a modified sensing circuit and its application in a hierarchical memory architecture is shown. The source terminal of the output transistor is now connected to a virtual ground line xv which enables the same signaling scheme as used for the

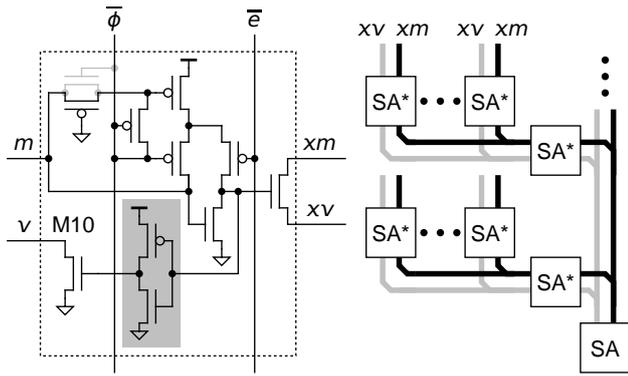


Fig. 9. The modified sensing circuit with a virtual ground on the output side and its implementation in a hierarchical memory architecture. SA* denotes the modified sensing circuit.

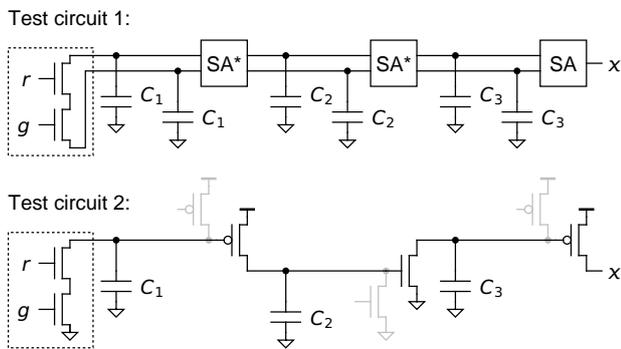


Fig. 10. Test circuits for sensing scheme comparison. Test circuit 1 shows the proposed sensing scheme with virtual grounds. Test circuit 2 shows a domino-style sensing scheme. The gray shaded transistors are used to pre-charge the digit-lines and are not relevant for the access time comparison. The dotted box encloses the read-only port of the (modified) 8T-SRAM cell.

modified 8T-SRAM cell. Since now the output signal xm will not fall below the threshold voltage of the virtual ground switch M10, the output signal xm cannot be used to control the gate of M10 anymore. Thus, it is required to generate the control signal for the virtual ground switch inside the sensing circuit with an additional inverter (highlighted in Fig. 9 by a gray shaded box).

To show the performance improvement of the proposed signaling scheme, two test circuits representing the data signal path of a hierarchical memory architecture were designed and compared by simulations. The first test circuit employs the proposed sensing circuits and pairs of digit line and virtual ground line on each level of hierarchy. The second test circuit employs a domino-style sensing circuit similar to the left circuit of Fig. 7. However, in domino-style sensing circuits successive stages alternatively employ NMOS and PMOS sensing devices. The test circuits are depicted in Fig. 10. For fair comparison, each corresponding bit-line,

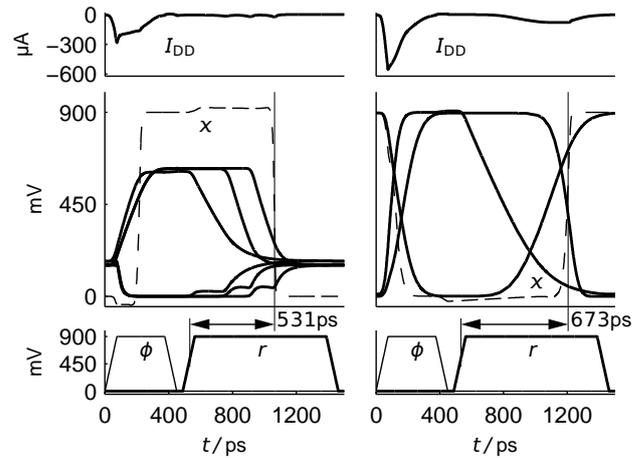


Fig. 11. Comparison of simulated waveforms for a read-cycle of the proposed sensing circuit on the left and a domino-style full-swing signaling circuit on the right (slow corner, 125°C, $V_{DD} = 900\text{mV}$).

digit-line, and virtual ground of the test circuits are equally loaded with parasitic capacitances on each hierarchy level. Figure 11 shows the simulated waveform for read-cycles of both sensing circuits. The left read-cycle was simulated for the proposed signaling scheme with included virtual ground concept. This waveform shows the same properties as the waveform in Fig. 3: the virtual grounds are successively released as soon as the data bit is detected. The reduced voltage swing of approximately 475mV leads to digit-line energy savings of approximately 50%. The right waveform shows the full-swing signals of the domino-style sensing circuits. The access time of the proposed sensing circuit ($t_{ACC} = 531\text{ps}$) is smaller than the access time of the domino-style sensing scheme ($t_{ACC} = 673\text{ps}$). This leads to an access time improvement of approximately 21%.

5 Conclusions

A new sensing circuit for a single-ended read-only-port of SRAM cells is introduced. Instead of pre-charging the bit-line to V_{DD} , the proposed circuit sets the pre-charge level close to the threshold level $V_{DD} - |V_{th}|$ of the individual sensing device, while ensuring a process variation tolerant bit-line noise margin. With a small modification of the 8T-SRAM cell to provide an additional bit-line v , the charge dissipation of the bit-line m during evaluation can automatically be stopped by the proposed circuit, once the data bit is detected by the sensing circuit. Both effects together reduce the bit-line swing and the bit-line power dissipation. For $N = 64 \dots 128$ cells, the proposed circuit achieves fewest energy compared to the reference circuits along with good performance improvements. The proposed sensing circuit is also suited for a hierarchical memory architecture utilizing the optimized pre-charge level and virtual ground concepts.

References

- Chang, L., Fried, D. M., Hergenrother, J., Sleight, J. W., Dennard, R. H., Montoye, R. K., Sekaric, L., McNab, S. J., Topol, A. W., Adams, C. D., Guarini, K. W., and Haensch, W.: Stable SRAM Cell Design for the 32 nm Node and Beyond, in: Symposium on VLSI Technology, 128–129, 2005.
- Chang, L., Nakamura, Y., Montoye, R. K., Sawada, J., Martin, A. K., Kinoshita, K., Gebara, F. H., Agarwal, K. B., Acharyya, D. J., Haensch, W., Hosokawa, K., and Jamsek, D.: A 5.3 GHz 8T-SRAM with Operation Down to 0.41 V in 65 nm CMOS, in: Symposium on VLSI Circuits, 252–253, 2007.
- Cosemans, S., Dehaene, W., and Catthoor, F.: A Low-Power Embedded SRAM for Wireless Applications, *IEEE Journal of Solid-State Circuits*, 42, 1607–1617, 2007.
- Qazi, M., Stawiasz, K., Chang, L., and Chandrakasan, A.: A 512 kb 8T SRAM Macro Operating Down to 0.57V with An AC-Coupled Sense Amplifier and Embedded Data-Retention-Voltage Sensor in 45 nm SOI CMOS, in: *IEEE International Solid-State Circuits Conference*, 350–351, 2010.
- Takeda, K., Hagihara, Y., Aimoto, Y., Nomura, M., Uchida, R., Nakazawa, Y., Hirota, Y., Yoshida, S., and Saito, T.: Per-Bit Sense Amplifier Scheme for 1GHz SRAM Macro in Sub-100 nm CMOS Technology, in: *IEEE International Solid-State Circuits Conference*, 1, 502–542, 2004.
- Verma, N. and Chandrakasan, A. P.: A High-Density 45 nm SRAM Using Small-Signal Non-Strobed Regenerative Sensing, *IEEE Journal of Solid-State Circuits*, 44, 163–173, 2009.