A High-Density 45 nm SRAM Using Small-Signal Non-Strobed Regenerative Sensing

Naveen Verma, *Student Member, IEEE*, and Anantha P. Chandrakasan, *Fellow, IEEE*

***Abstract*—High-density SRAMs utilize aggressively small bit-cells, which are subject to extreme variability, degrading their read SNM and read-current. Additionally, array performance is also limited by sense-amplifier offset and strobe-timing uncertainty. This paper, presents a sense-amplifier that targets all of these performance degradations: specifically, simple offset compensation reduces sensitivity to variation while imposing minimal loading on high-speed nodes; stable internal voltage references serve as an internal means to self-trigger regeneration to avoid tracking mismatch in an external strobe-path; precise small-signal detection withstands small read-currents so that other bit-cell parameters can be optimized; and single-ended sensing provides compatibility to asymmetric bit-cells, which can have improved operating margins. The design is integrated with a 64-kb high-density array composed of 0.25 μm^2 6T bit-cells. A prototype, in low-power 45 nm CMOS, compares its performance with a conventional sense-amplifier, demonstrating an improvement of 4X in access-time sigma and 34% in overall worst case access time.**

***Index Terms*—Auto-zeroing, device variation, offset compensation, sense-amplifier, SRAM.**

I. Introduction

HIGH-DENSITY SRAMs are a critical means for digital systems to benefit from technology scaling. Further, more and more, these systems target highly mobile applications, where low power consumption is paramount. As a result, their constituent SRAMs, whose power is becoming increasingly dominant, must utilize low-power (LP) CMOS technologies.

At the 45 nm node, bit-cells as small as a 0.25 μm^2 are manufacturable. Fig. 1, however, summarizes the associated challenges. First, such aggressive scaling requires the use of tiny devices, aggravating variability. As a result, the cell read current and read static noise margin (SNM) [1], [2], are greatly degraded, subjecting the 6T design to a highly constrained density–stability–performance trade-off. This is, however, alleviated somewhat in 8T topologies at the cost of density and sensing complexity. Similarly, sense-amplifiers are also primarily limited by variation, as their input offset sets the required bit-line discharge, and, as a result, in many cases, their area has stopped scaling [3]. Accordingly, even though their density constraints are not nearly as severe as the bit-cell's, sense-amplifiers are becoming a limiting feature in high-density arrays. Finally, another source of variation emerging as a major limitation is the timing of the sense-amplifier strobe signal, which does not track well with the array read-path over all of the operating corners. As a result, excessive margin must be introduced into the critical read-path.

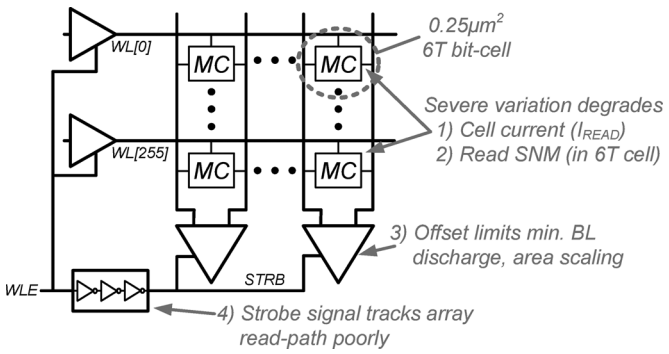

Fig. 1. Critical limitations faced by low-power high-density SRAMs at the 45 nm node.

Although these limitations span the entire array, they can be greatly alleviated by means of the sense-amplifier. The following sections start by elaborating on the challenges faced by low-power high-density SRAMs and how they are related to the sense-amplifier. Then, the emerging need for efficient single-ended sensing schemes are developed. Finally a new sense-amplifier structure is introduced to address the challenges, and prototype results are presented.

II. High-Density SRAM Design Challenges

The most urgent challenges in high-density SRAMs arise from severe variation in the bit-cell and the trade-offs inherent in its structure. The following subsections describe how these are related to the sense-amplifier.

A. Read-Current Degradation

Low-power SRAMs must incur the reduced I_{READ} that comes with technology optimizations to manage leakage-currents. Fig. 2(a) shows how read-current scales further with cell size. The reduction in mean read-current is a direct consequence of reducing the size of the driver devices. However, the increased variation in the smaller devices also results in severer degradation to the weak-cell read-current in proportion to this already reduced mean read-current. For instance, the

Manuscript received April 14, 2008; revised June 18, 2008. Current version published December 24, 2008. IC fabrication was provided by Texas Instruments. The work of N. Verma is supported by the Intel Foundation Ph.D. Fellowship Program and NSERC. Partial support provided by Defense Advanced Research Projects Agency (DARPA).

The authors are with Microsystems Technology Laboratories, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139 USA (e-mail: nverma@mtl.mit.edu; anantha@mtl.mit.edu).

Digital Object Identifier 10.1109/JSSC.2008.2006428

0018-9200/$25.00 © 2008 IEEE

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

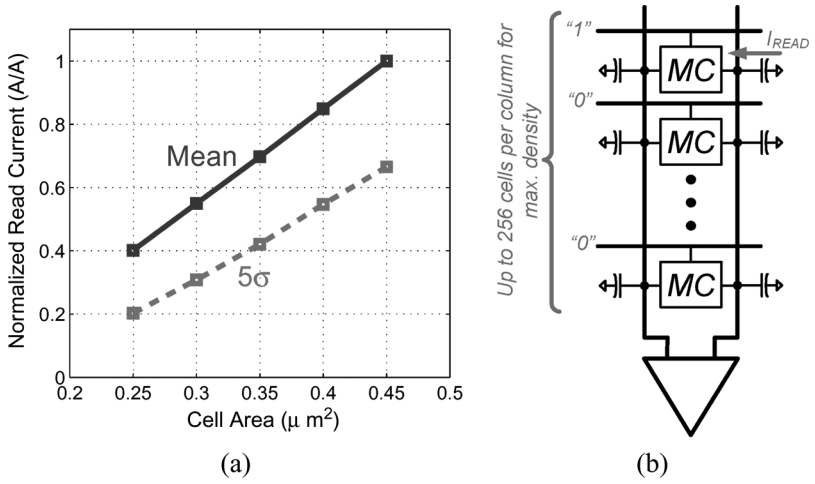

Fig. 2. Degradation in bit-line discharge time for high-density SRAMs caused by (a) reduced cell read-current and (b) increased bit-line capacitance.

5σ read-current for the 0.25 μm^2 cell is easily degraded by an additional factor of two.

Furthermore, in high-density arrays, the integration of many cells per bit-line, in order to maximize array efficiency, leads to very large bit-line capacitance. As shown in Fig. 2(b), the resulting ratio of I_{READ}/C_{BL}, which is critical to the array's performance, suffers even further.

B. Read SNM Degradation

The increased variation in the highest density bit-cells degrades the read SNM in the same manner as the read-current [2]. Fig. 3(a) shows this reduction in the presence of variation (at the 5σ level) as the cell area is scaled. Accordingly, in high-density bit-cells, read-current and read SNM, which are both critical metrics, are simultaneously stressed. Exacerbating the matter even further, in 6T bit-cells, these exhibit a strong inverse correlation, as shown in Fig. 3(b) [4]. This comes about as a result of the bit-cell access-devices, which must be strong for maximum read-current but weak for maximum read SNM. Unfortunately, then, optimizations targeting one are likely to worsen the other. Nonetheless, for functionality, read-stability remains the paramount concern, and cell design trends and circuit assists preferentially aim to improve read SNM, leading to further reduction in read-current [5].

Accordingly, density and stability trade-offs imply the urgent need for an alternate method to recover the ensuing loss in array performance. Although 8T bit-cells overcome this interdependence, as mentioned in Section III, they impose additional requirements with regards to data sensing.

C. Sense-Amplifier Strobing Uncertainty

The sense-amplifier strobe signal must ensure that enough time is allocated for bit-line discharge to overcome the sense-amplifier offset. However, as shown in Fig. 4(a), the bit-line discharge time depends on the delay through the array read-path, which is limited by the weakest bit-cell. Such a cell, whose statistical characteristics might be beyond the 5σ level, is impossible to replicate in the strobe control path. Consequently, as shown in Fig. 4(b), the timing of the two paths diverges greatly over process, voltage, and temperature corners [6] even if carefully laid out SRAM devices are employed in the strobe path, as in [7].

Since the strobe path must be designed such that it is longer than the array path in all cases, this implies, as shown in Fig. 4(b), that in many cases, it will be much longer than it needs to be, thereby excessively limiting the overall performance. In fact, the overall worst case delay, for the array configuration considered, need only be 820 ps based on the read-path delay; however, the actual worst case delay is 980 ps, limited by the strobe-path, imposing an excess overhead of nearly 20%.

III. Single-Ended Sensing

One way to address the severe stability limitation plaguing the bit-cell is to use an alternate structure. Asymmetric cells, both 7T [4] and 8T [8], can have much wider operating margins, making them particularly compelling in the face of increasing variation [9]–[11]. However, neither of these provides a differential read-port, and, therefore, they require a compatible means for efficient single-ended sensing.

The most common technique used with 8T designs is full-swing sensing [3]. Here, the read bit-line must discharge almost completely so that the read data can be detected by a sensing structure that may be reduced to essentially a logic gate. To maintain performance, however, bit-line discharge times must be minimized through the use of large bit-cells, to increase read-current, and very short bit-lines, to reduce bit-line capacitance. Unfortunately, both of these directions severely degrade the density.

Alternatively, small-signal sensing can be retained by employing a pseudo-differential sense-amplifier [12]. However, here, generation of the complementary reference is unavoidable, and, importantly, it must track all operating conditions even with respect to cells in the tails of the arrays statistical distribution. As a result, its design and testability is highly complicated.

A. Noise Sensitivity

An important drawback to any single-ended sensing scheme is the loss of common-mode noise rejection capability both on

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

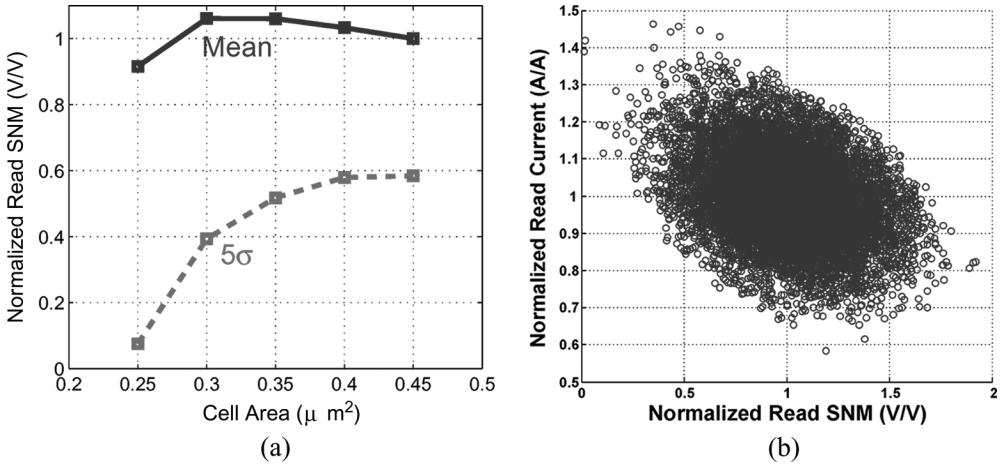

Fig. 3. Read SNM trade-off in high-density SRAMs limited by (a) cell size and (b) inverse correlation with cell read-current, caused by opposing access-device requirements.

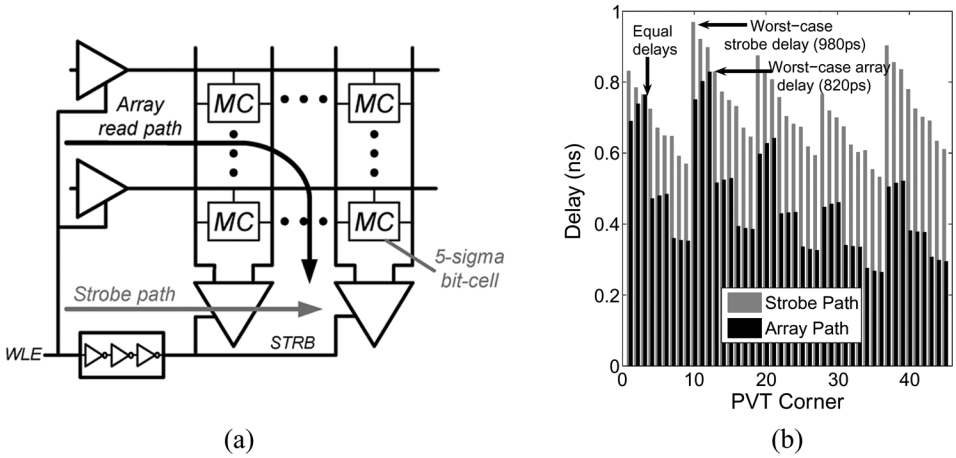

Fig. 4. Array read-path and sense-amplifier strobe-path (a) limited by matching to 5σ bit-cell and (b) exhibiting severe delay divergence over process-voltage-temperature conditions, leading to excess overall delay.

the power-supply and on the bit-lines. As a result, an inherent trade-off is introduced between sensitivity, which ultimately affects the array's performance, and robustness to noise. Since no distinction can be made between noise on the bit-line and read-data droop, it becomes critical to ensure a desired level of bit-line noise margin, and relatedly, the sensitivity (or, correspondingly, speed) must be characterized with respect to this noise margin.

IV. Non-Strobed Regenerative Sense-Amplifier

The Non-Strobed Regenerative Sense-Amplifier (NSR-SA) shown in Fig. 5 addresses the limitations described in Section II. Two cascaded inverters ($M1$–$M2$ and $M3$–$M4$) form an amplification path that is self-biased for high-gain by the feedback switches, S_{AZ}. As described in Section IV-B, these also perform offset compensation of the inverter amplifiers via input auto-zeroing [13]. Importantly, however, to support small-signal sensing, which greatly improves the density-performance trade-off of the array, the large gain required is achieved most efficiently through regeneration [14]. Accordingly, the regenerative device, $M5$, provides very large positive-feedback gain. A critical feature, however, is that regeneration requires no external enable or strobe signal, thereby overcoming the severe tracking uncertainty described in Section II-C. Instead, the ideal DC transfer function, shown in Fig. 5, is implemented in continuous time, and the precise point at which regeneration is enabled depends only on the input bit-line voltage (with respect to the internal reference, V_{TRIP}). Further, V_{TRIP} is very stable despite variation since it is generated implicitly by the original auto-zeroing, and, additionally, its precise value can be selected by design. Accordingly, this sense-amplifier, which is single-ended and compatible with asymmetric bit-cells, can enforce a desired balance between sensitivity and noise rejection.

A. Basic Operation

The NSR-SA operates over two phases: reset and detection. The reset phase occurs during SRAM bit-line pre-charge, during which the sense-amplifier does not need to detect bit-line droop. Hence, this time is used to perform self-correction of offsets via auto-zeroing. The detection phase corresponds with bit-line discharge, where the actual read-data must be quickly resolved.

1) Reset Phase: The purpose of the reset phase is to charge the internal nodes so that the inverters formed by $M1$–2 and

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

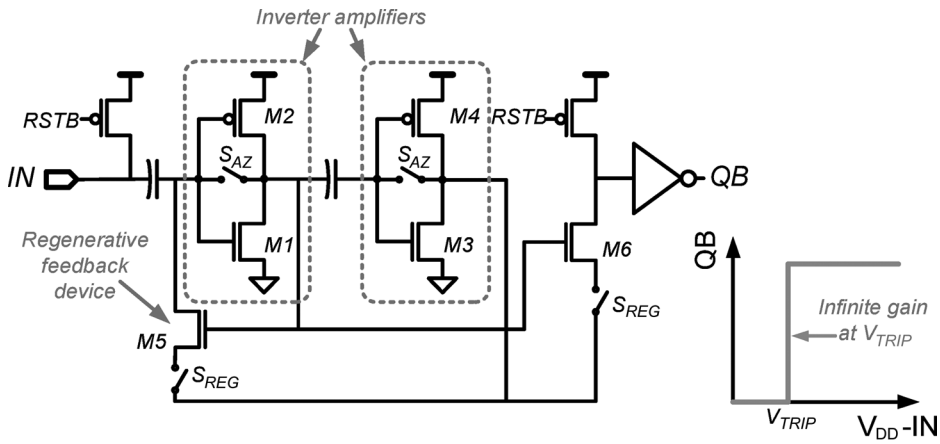

Fig. 5. Non-strobed regenerative sense-amplifier (NSR-SA) schematic and ideal transfer function.

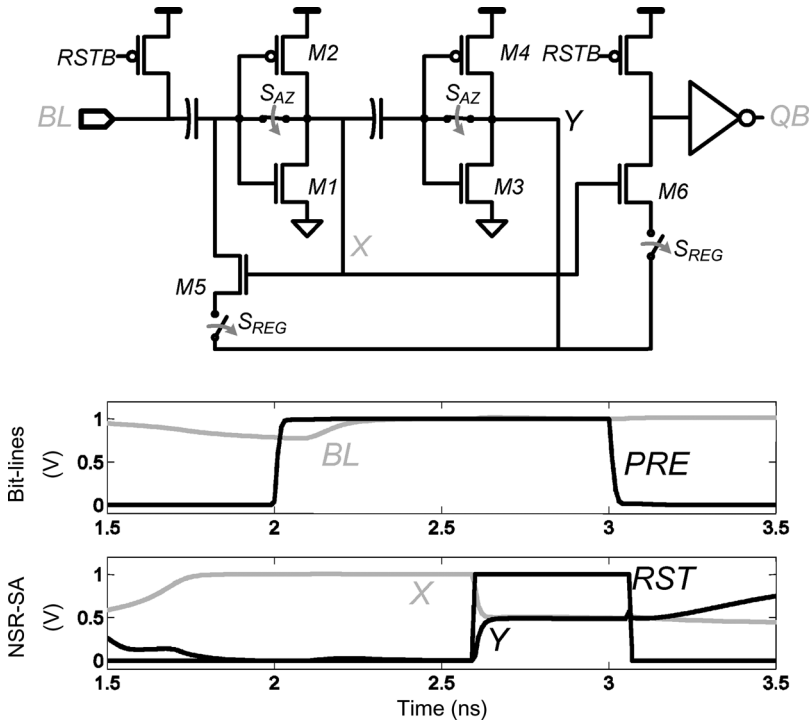

Fig. 6. NSR-SA circuit and waveforms during reset phase.

$M3$–4 are biased in their high-gain regions very close to their ideal trip-points. Simultaneously, this initializes the regenerative device, $M5$, such that its positive feedback gain is very low.

As mentioned, the reset phase occurs during bit-line pre-charge; however, it is actually meant to occupy only a small portion of this period. For instance, as shown in Fig. 6, a short RST pulse is asserted (even its duration as shown is much longer than required). During this time, the input node and output stage $(M6)$ are pre-charged, while, simultaneously, the negative feedback switches, S_{AZ}, are closed, and the regeneration switches, S_{REG}, are opened. Consequently, as shown in Fig. 6, nodes X and Y get biased to mid-rail voltages at the inverter trip-points, which are nominally designed to be equal. Offsets can lead to differences in their precise value, but the analysis in Section IV-B describes how this biasing greatly attenuates the errors.

It can be seen that nodes X and Y settle to their required reset values in less than 100 ps, implying that only a very small RST pulse is required. It should be noted, however, that while nodes X and Y remain at these mid-rail voltages, a static current path exists through $M1$–2 and $M3$–4. As discussed in Section V-B, however, the total power overhead introduced as a result is small.

2) Detection Phase: Following reset, bit-line discharge must be detected. First, the case where the bit-line remains high at its pre-charge value (i.e. logic "1") will be considered. Here, the bit-line voltage, BL, remains unchanged, so all internal voltages remain essentially at their reset bias values. For instance, as shown in Fig. 7, after the negative feedback switches, S_{AZ}, are opened, nodes X and Y remain unchanged aside from small perturbations arising from charge-injection errors from the S_{AZ} switches (which will be addressed in Section IV-C). Accord-

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

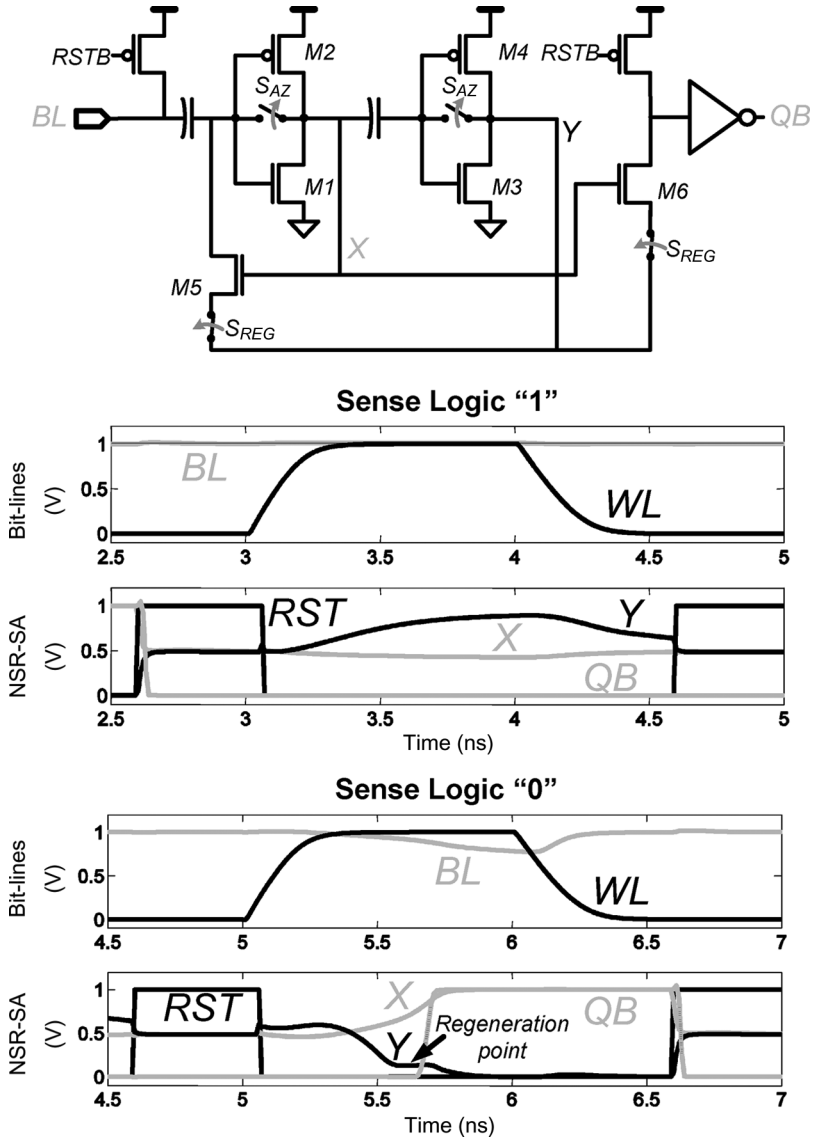

Fig. 7. NSR-SA circuit and waveforms during detection phase (for both bit-line logic cases).

ingly, when the S_{REG} switches are closed to enable $M5$ and $M6$, the V_{GS} of these devices remains very small, and in fact, as discussed in Section IV-C, even negative, thanks to charge-injection errors originating from the specific choice of the devices used to implement the S_{AZ} switches. As a result, the output logic level of $QB = 0$ is sustained. Of course, bit-line leakage, arising from the unaccessed cells sharing BL, can compromise the logic "1" value, leading to detection errors. Although, the NSR-SA's BL noise margin can be designed to reject these (as discussed in section Section IV-D), maximum-leakage simulations in the target LP process indicate that, for the 256 cells/BL configuration considered, the nominal NSR-SA rejects bit-line leakage for access times up to 35 ns, which is much longer than the target delays.

Alternatively, in the logic "0" case, also shown in Fig. 7, an intentional bit-line droop is detected. Here, the voltage of node X rises rapidly, as a result of the inverter gain, and the voltage of node Y decreases even more rapidly, as a result of the cascaded inverter gains. Correspondingly, at some point, the regenerative device's (i.e., $M5$'s) V_{GS}, which is the difference between the voltage of nodes X and Y, becomes large enough that the device turns on, triggering positive feedback. Subsequently, the input of the first inverter is actively pulled low, and the entire NSR-SA quickly latches. The precise point where regeneration is enabled can be seen at the annotated inflection in the waveform of node Y. Shortly after this, the output, QB, quickly changes its state.

3) Output Clocking Margin: As mentioned, an important feature of the NSR-SA is that regeneration is triggered by the input bit-line droop itself, rather than an explicit strobe signal. Nonetheless, the read data must ultimately be clocked at the array output. However, as shown in Fig. 8(a), the timing problem described in Section II-C has actually been overcome and not just propagated to the subsequent output clock, $OCLK$.

By comparison, for instance, the strobe signal of a conventional sense-amplifier must arrive during the bit-line discharge phase, before the next pre-charge phase can begin. As a result, the timing of the strobe signal limits the critical read-path inside the array. However, as shown in Fig. 8(b), the NSR-SA latches

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

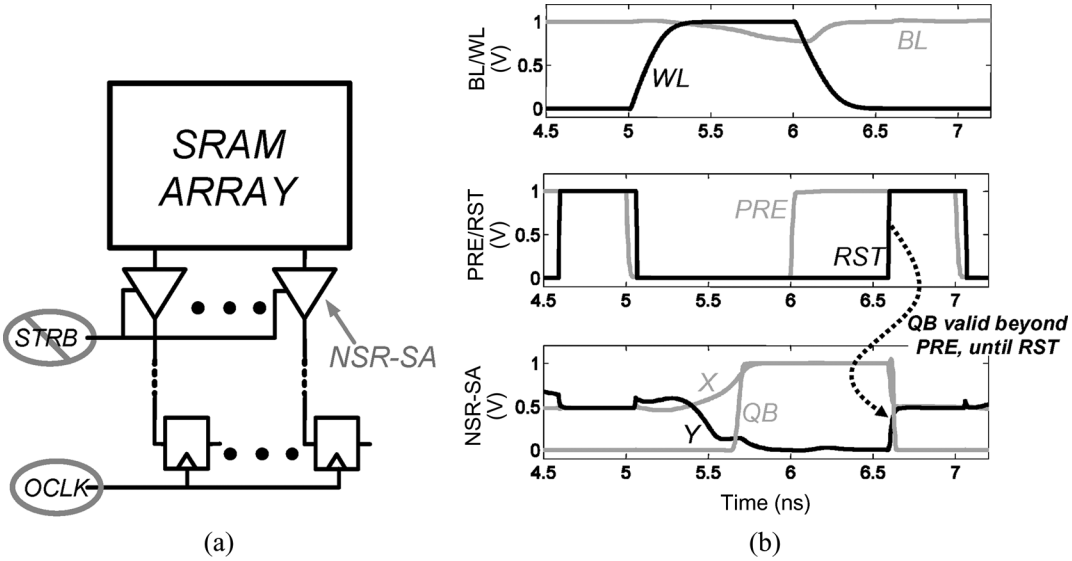

Fig. 8. Output clocking (a) at array-level with (b) waveforms showing decoupling from internal critical read-path.

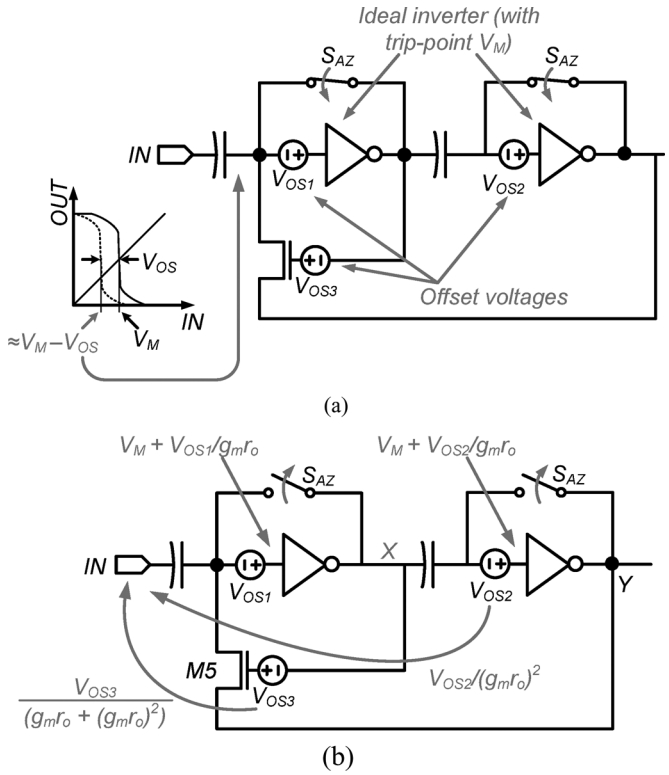

Fig. 9. Offset compensation (a) technique and (b) analysis.

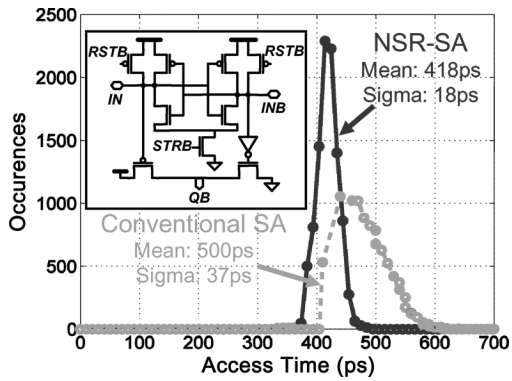

Fig. 10. 10 k point Monte Carlo simulation showing improved sigma of NSR-SA access time compared to conventional sense-amplifier access time.

the output without a strobe signal, and the output state remains valid until the reset pulse is asserted. Since the reset pulse can be a small fraction of the pre-charge phase, the output data can be clocked well after the pre-charge phase begins, until the data is cleared by the reset pulse. As a result, its timing is decoupled from the critical read-path inside the array.

B. Offset Compensation

In Fig. 9(a), the inverters are abstracted by their logic symbol, and each of them is modeled as being ideal and offset-free. The offsets are modeled as the series voltage sources, V_{OS1-3}, each of which is associated with an inverter and the regenerative device.

During the reset phase, the S_{AZ} switches are closed, forcing the inverters into negative feedback. Analytically, this implies that the input and output values of the voltage transfer characteristic (VTC) must be equal, as shown by the diagonal line. Accordingly, if the inverter is offset-free, as in the case of the solid VTC, its input would settle to the voltage V_M, which is the ideal trip-point. However, offset shifts the VTC to the left by an amount equal to V_{OS}. Now, as shown by the dotted VTC, the input instead settles to a value approximately equal to $V_M - V_{OS}$. In particular, this requires that the VTC is nearly vertical near these trip-points, which corresponds to a large inverter gain. If this is the case, the resulting voltage of $V_M - V_{OS}$ gets stored on the preceding coupling capacitor. As a result, $-V_{OS}$ appears in series with the actual V_{OS}, canceling the offset-voltage and biasing each inverter to its ideal trip-point.

Complete offset cancellation in this manner, however, requires that the gains of the inverters be infinite. Practically, the offset is only reduced by a factor equal to the finite inverter gain [13], which is approximately given by $g_m r_o$, where g_m

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

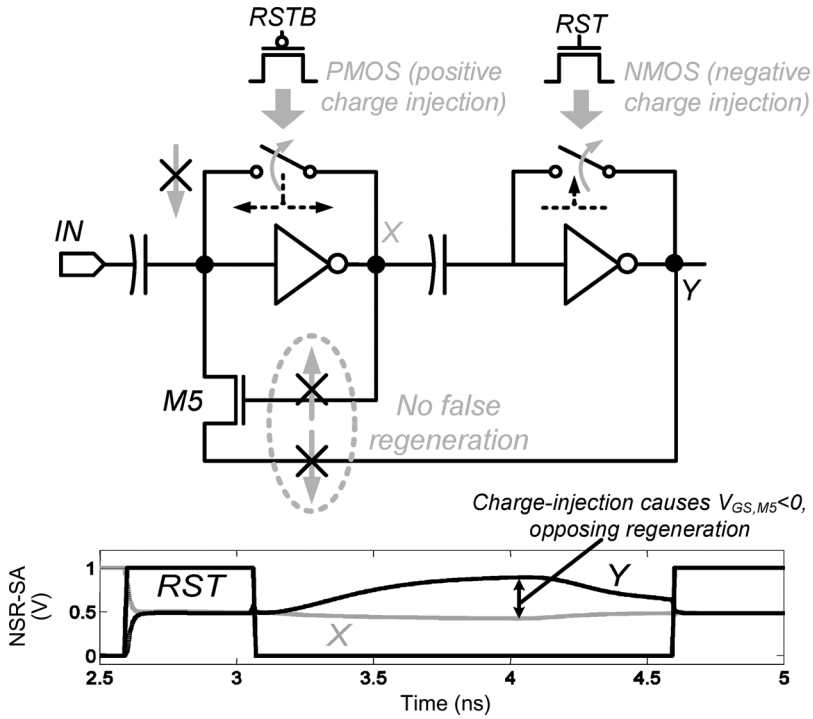

Fig. 11. NSR-SA robustness to false-regeneration in the presence of charge-injection errors.

is the transconductance and r_o is the output resistance of the inverter devices. As shown in Fig. 9(b), however, the residual offset of the second inverter stage is reduced by an additional factor of $g_m r_o$ when input referred, which requires dividing by the preceding inverter's gain. Finally, nothing explicit is done to manage the offset of the regenerative device, but, once again, its contribution to the overall offset is very small when input referred, since the gain from the input to its V_{GS} is given by $g_m r_o + (g_m r_o)^2$ through the inverter path. Accordingly, after all of the attenuation factors in Fig. 9(b), the offset of the first inverter dominates, but it is significantly reduced thanks to the offset compensation.

A Monte Carlo simulation of the access-time distributions for the NSR-SA and a conventional strobed sense-amplifier is shown in Fig. 10. In this context, access time includes the delay of the word-line driver, bit-line discharge, and sense-amplifier, and it is measured from the word-line enable signal to the sense-amplifier output valid. A 256-by-256 array configuration is considered with a mean bit-cell, so that the variation in the sense-amplifiers can be isolated. The conventional sense-amplifier has been sized for minimum offset while occupying a layout area of 12 μm^2, and it also includes the strobe timing margin required to ensure sufficient bit-line discharge in all operating conditions. Nonetheless, as shown, it achieves good mean performance thanks to its differential operation, as, nominally, it must wait long enough for only a very minute differential to develop on the bit-lines. However, because it is more sensitive to variation, its sigma is much worse at 37 ps, compared to 18 ps for the NSR-SA, and, therefore, it achieves far slower worst case performance.

The greatly improved stability achieved by the NSR-SA points to the benefit of offset compensation. However, the next primary limitation to its sigma comes from variation in the charge-injection errors introduced by the S_{AZ} switches.

C. False Regeneration Immunity

An even more urgent failure mode arising from charge-injection errors than residual delay sigma, however, is that, in an amplifier with high gain and regeneration, these can potentially result in resolution to the wrong output state, from which recovery might not even be possible. For instance, in the NSR-SA, if charge-injection errors were to cause the input of the first inverter to decrease and the input of the second inverter to increase, the gain through the inverters would cause a large positive V_{GS} on the regenerative device, $M5$, causing the output to transition and latch, regardless of the input bit-line voltage.

The NSR-SA, however exploits the fact that it only needs to respond to bit-line discharge, not up-charge. As a result, regeneration only needs to occur in one direction, and, as shown in Fig. 11 the charge-injection error sources can be designed to oppose that direction. Specifically, the main device used for the first S_{AZ} switch is a PMOS, whose charge-injection errors tend to increase the input of the first inverter, and the main device used for the second S_{AZ} switch is an NMOS, whose charge-injection errors tend to decrease the input of the second inverter. As a result, nodes X and Y decrease and increase respectively, giving a negative V_{GS} on $M5$, pushing it away from false regeneration.

D. Regeneration Trip Point

Although offset compensation improves the stability of the NSR-SA in the presence of variation, it is also important to set its nominal trip-point, V_{TRIP}, based on speed and noise rejection considerations. One way to achieve this control is by

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

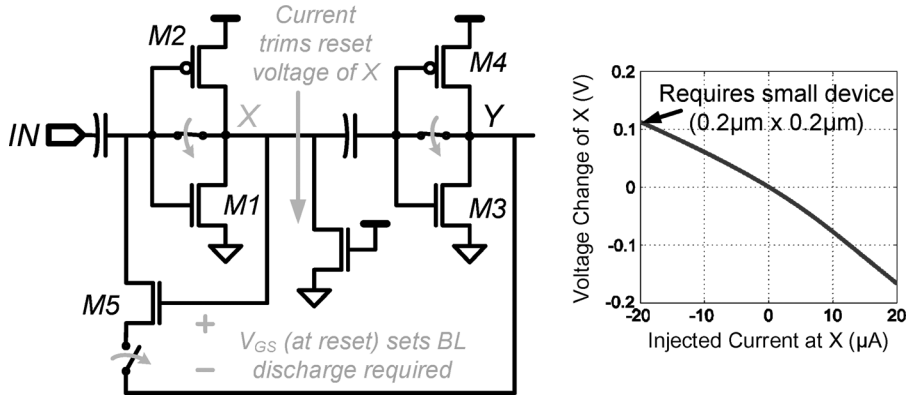

Fig. 12. NSR-SA technique to set regeneration trip-point (V_{TRIP}) for noise rejection and sensitivity considerations.

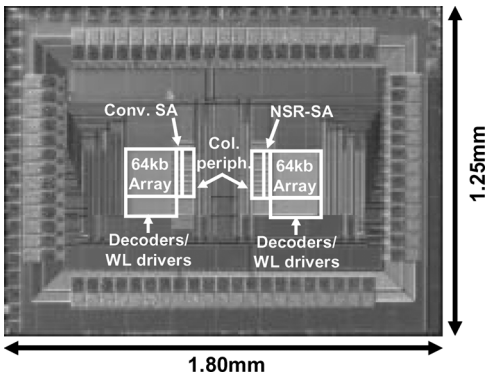

Fig. 13. IC die photo of prototype implemented in low-power 45 nm CMOS to compare performance of NSR-SA with conventional sense-amplifier.

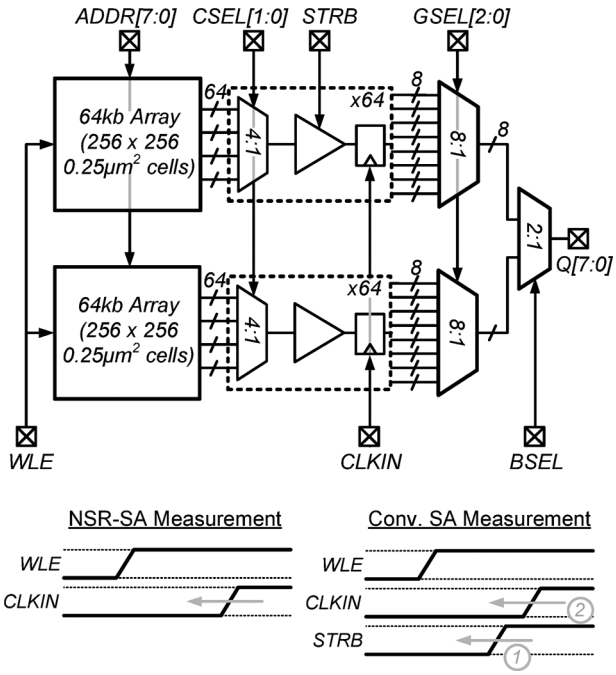

Fig. 14. Block diagram of prototype test-chip and access-time measurement methodology.

adjusting the reset voltages of nodes X and Y, which change the V_{GS} of $M5$ and trim the amount of additional bit-line discharge required to actually trigger regeneration. This can be implemented, for instance, with the addition of an appropriately sized device at the inverter outputs, as shown in Fig. 12. This device can be either an NMOS or PMOS (NMOS is shown), depending on the trimming required, and, with its gate connected to an appropriate rail voltage, even a very small device (less than $0.2\ \mu\text{m} \times 0.2\ \mu\text{m}$) results in a wide range of output reset voltages (covering almost 0.3 V, in this case).

V. Test-Chip

A photograph of the prototype test-chip, which is implemented in LP 45 nm CMOS is shown in Fig. 13. A block diagram of its architecture is shown in Fig. 14. Here, two 64-kb (256 by 256) arrays of high-density 0.25 μm^2 6T SRAM bit-cells are integrated. The first drives a set of conventional strobed sense-amplifiers, of the structure shown in Fig. 10, and the second drives a set of NSR-SAs, so that the relative performances can be accurately compared.

To measure the access time with the NSR-SA, the word-line enable signal, WLE, must be asserted, and the $CLKIN$ signal must be swept inward, as shown in Fig. 14, until bit-failures appear. To measure the access time with the conventional sense-amplifier, however, first the $STRB$ signal must be swept inward, with the $CLKIN$ signal held at a much larger delay, and then the $CLKIN$ signal must be swept in to measure the final access time. This two sweep procedure is also shown in Fig. 14. The final $WLE - CLKIN$ access times can then be suitably compared, since the WLE and $CLKIN$ paths for the two arrays are carefully matched on-chip.

A. Noise Characterization

As mentioned in Section III-A, with any single-ended sensing scheme, it is important to characterize the noise rejection versus sensitivity performance. In the case of bit-line noise, this requires dedicated additional circuitry, which is shown in Fig. 15, to inject a controllable noise amplitude on one set of bit-lines. In particular, the coupling capacitors, C_{NOISE} can be pulsed from off-chip, and the bit-line noise can be estimated from the amplitude of the off-chip pulse (set by the potentiometer), and the nominal ratio of the capacitor divider formed by C_{NOISE} and the actual parasitic bit-line capacitance, C_{BL}. Simultaneously, the regeneration trip-point of the NSR-SA can be adjusted by injecting a controllable current at the inverter outputs using $M7$

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

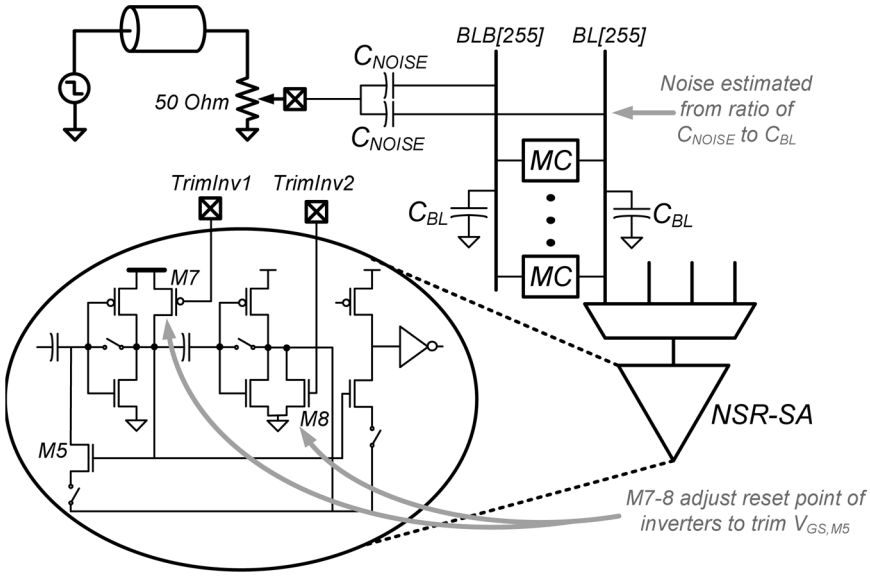

Fig. 15. Dedicated circuitry to inject a controllable noise amplitude on one set of bit-lines and independently adjust the sensitivity/noise rejection of the NSR-SA.

and $M8$, whose gates are biased off-chip. As a result, this allows both the bit-line noise and NSR-SA sensitivity to be varied, allowing noise rejection to be characterized versus access-speed.

B. Measurement Results

Measurements were taken from 53 chips in order to capture some statistical behavior. The access-time distributions for the NSR-SA and the conventional sense-amplifier are shown in Fig. 16. Additionally, a distribution of the difference between the access times on each chip is also plotted to de-embed the absolute delays through board traces and chip packaging. As shown, the NSR-SA achieves superior sigma, as expected from Monte Carlo simulations. Unlike Fig. 10, however, the measurement results include the effect of variation in the bit-cells, and as a result the absolute delays here are larger. Nonetheless, especially in the presence of this extra variation, which implies lower worst case bit-cell read-current, the NSR-SA performs very well, achieving a factor of four reduction in access-time sigma. Accordingly, the overall worst case delay improves from 2.46 ns to 1.63 ns, representing a speed-up of 34%.

The bit-line noise rejection observed with respect to access time is shown in Fig. 17. Here, the NSR-SA is tuned for each plotted access-time point, and the bit-line noise is increased until erroneous data is observed, yielding the corresponding bit-line noise rejection point. In the case shown, for instance, the NSR-SA can be tuned for nearly 50 mV of noise margin, corresponding to an increased amount of bit-line discharge required for data sensing. Alternatively, the speed can be increased by tuning for increased bit-line discharge sensitivity at the cost of noise margin.

With regards to transient supply noise, no severe effects in functionality were observed with large negative pulses on V_{DD}. However, long positive pulses larger than 55 mV that occur after the reset phase must be managed with proper system timing in order to avoid incorrect data latching.

A performance summary of the NSR-SA and conventional sense-amplifier is provided in Table I. Each conventional sense-amplifier occupies a layout area of 12 μm^2 while each NSR-SA occupies 19 μm^2, though this includes all of the testability features that have been integrated for characterizations purposes

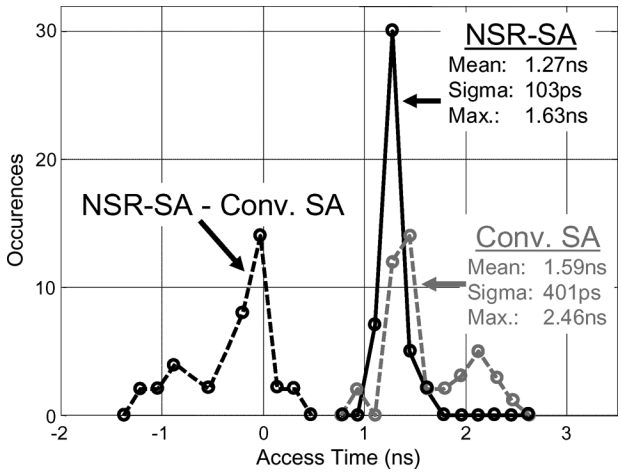

Fig. 16. Access-time measurements from 53 chips showing a factor of four improvement in the NSR-SA distribution sigma compared to the conventional sense-amplifier sigma.

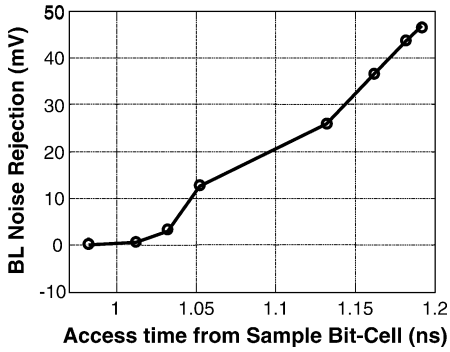

Fig. 17. Measured bit-line noise rejection with respect to access time, showing ability to tune one at the cost of the other.

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

TABLE I
TEST-CHIP PERFORMANCE SUMMARY

	Conventional SA	NSR-SA
Technology	45nm low-power CMOS	
Cell size	0.25μm²	
Array configuration	256 x 256	
Capacity	64kb	
Area	12μm²	19μm²*
Power in reset	-	23μW
% of array power (at 100MHz)	2%	7%
Mean access-time	1.59ns	1.27ns
Max. access-time	2.46ns	1.67ns
Access-time sigma	401ps	103ps

*Includes testability features

only. Additionally, during the reset phase, each NSR-SA draws static power of 23 μW while its inverters are at their trip-points, and, in total, this contributes 7% to the overall array power when the array is running at 100 MHz.

VI. CONCLUSIONS

Low-power, high-density SRAMs are critical to technology scaling in a broad range of applications and devices. Their reliability, however, is threatened by extreme variability, and excessive margining to cell stability, sense-amplifier offset, and array timing is required. As a result, severe performance compromises must be incurred. To recover these, however, an offset compensating self-regenerating sense-amplifier is integrated into a 64-kb array of high-density 0.25 μm^2 bit-cells in low-power 45 nm CMOS. By performing simple auto-zeroing to detect and correct its own offsets, the sense-amplifier achieves high-stability in the presence of variation. Further, the stable internal voltage references are exploited as a means to self-trigger regeneration without relying on an external strobe-path, thereby precluding the associated timing uncertainty. The ability to stably detect small-signal bit-line discharge implies that smaller read-currents can be accommodated for a given performance requirement, and the bit-cells can thus be optimized for stability, facilitating increased density.

ACKNOWLEDGMENT

The authors thank M. Clinton, X. Deng, T. Houston, and W.-K. Loh (TI) for their support and feedback.

REFERENCES

[1] E. Seevinck, F. J. List, and J. Lohstroh, "Static-noise margin analysis of MOS SRAM cells," *IEEE J. Solid-State Circuits*, vol. SC-22, no. 5, pp. 748–754, Oct. 1987.

[2] A. Bhavnagarwala, X. Tang, and J. Meindl, "The impact of intrinsic device fluctuations on CMOS SRAM cell stability," *IEEE J. Solid-State Circuits*, vol. 36, no. 4, pp. 658–665, Apr. 2001.

[3] K. Zhang, K. Hose, V. De, and B. Senyk, "The scaling of data sensing schemes for high speed cache design in sub-0.18 μm technologies," in *Symp. VLSI Circuits Dig.*, Jun. 2000, pp. 226–227.

[4] K. Takeda, Y. Hagihara, Y. Aimoto, M. Nomura, Y. Nakazawa, T. Ishii, and H. Kobatake, "The impact of intrinsic device fluctuations on CMOS SRAM cell stability," *IEEE J.Solid-State Circuits*, vol. 41, no. 1, pp. 113–121, Jan. 2006.

[5] M. Yabuuchi, K. Nii, Y. Tsukamoto, S. Ohbayashi, S. Imaoka, H. Makino, Y. Yamagami, S. Ishikura, T. Terano, T. Oashi, K. Hashimoto, A. Sebe, G. Okazaki, K. Satomi, H. Akamatsu, and H. Shinohara, "A 45 nm low-standby-power embedded SRAM with improved immunity against process and temperature variations," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2007, pp. 326–327.

[6] K. Sohn, N. Cho, H. Kim, K. Kim, H.-S. Mo, Y.-H. Suh, H.-G. Byun, and H.-J. Yoo, "An autonomous SRAM with on-chip sensors in a 80 nm double stacked cell technology," in *Symp. VLSI Circuits Dig.*, Jun. 2005, pp. 232–235.

[7] K. Osada, J.-U. Shin, M. Khan, Y.-D. Liou, K. Wang, K. Shoji, K. Kuroda, S. Ikeda, and K. Ishibashi, "Universal-Vdd 0.65–2.0 V 32 kB cache using voltage-adapted timing-generation scheme and a lithographical-symmetric cell," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2001, pp. 168–169.

[8] L. Chang, D. M. Fried, J. Hergenrother, J. W. Sleight, R. H. Dennard, R. Montoye, L. Sekaric, S. J. McNab, A. W. Topol, C. D. Adams, K. W. Guarini, and W. Haensch, "Stable SRAM cell design for the 32 nm node and beyond," in *Symp. VLSI Circuits Dig.*, Jun. 2005, pp. 128–129.

[9] L. Chang, Y. Nakamura, R. K. Montoye, J. Sawada, A. K. Martin, K. Kinoshita, F. H. Gebara, K. B. Agarwal, D. J. Acharyya, W. Haensch, K. Hosokawa, and D. Jamsek, "A 5.3 GHz 8 T-SRAM with operation down to 0.41 V in 65 nm CMOS," in *Symp. VLSI Circuits Dig.*, Jun. 2007, pp. 252–253.

[10] R. Joshi, R. Houle, K. Batson, D. Rodko, P. Patel, W. Huott, R. Franch, Y. Chan, D. Plass, S. Wilson, and P. Wang, "6.6+ GHz low Vmin, read and half select disturb-free 1.2 Mb SRAM," in *Symp. VLSI Circuits Dig.*, Jun. 2007, pp. 250–251.

[11] Y. Morita, H. Fujiwara, H. Noguchi, Y. Iguchi, K. Nii, H. Kawaguchi, and M. Yoshimoto, "An area-conscious low-voltage-oriented 8T-SRAM design under DVS environment," in *Symp. VLSI Circuits Dig.*, Jun. 2007, pp. 256–257.

[12] N. Tzartzanis and W. W. Walker, "A differential current-mode sensing method for high-noise-immunity, single-ended register files," in *IEEE ISSCC Dig. Tech. Papers*, Feb. 2004, pp. 506–507.

[13] C. C. Enz and G. C. Temes, "Circuit techniques for reducing the effects of op-amp imperfections: Autozeroing, correlated double sampling, and chopper stabilization," *Proc. IEEE*, vol. 84, no. 11, pp. 1584–1614, Nov. 1996.

[14] J.-T. Wu and B. A. Wooley, "A 100 MHz pipelined CMOS comparator," *IEEE J. Solid-State Circuits*, vol. SSC-23, no. 6, pp. 1379–1385, Dec. 1988.

Naveen Verma (S'04) received the B.A.Sc. degree in electrical and computer engineering from the University of British Columbia, Vancouver, BC, Canada, in 2003, and the M.S. degree from the Massachusetts Institute of Technology (MIT), Cambridge, MA, in 2005. He is currently pursuing the Ph.D. degree at the MIT, where his research interests include low-power mixed-signal circuits in the areas of analog-to-digital converters, SRAMs, and implantable biological systems.

Mr. Verma is the recipient of the Intel Foundation Ph.D. fellowship and the NSERC Postgraduate fellowship.

Anantha P. Chandrakasan (F'04) received the B.S, M.S., and Ph.D. degrees in electrical engineering and computer sciences from the University of California at Berkeley in 1989, 1990, and 1994, respectively.

Since September 1994, he has been with the Massachusetts Institute of Technology, Cambridge, MA, where he is currently the Joseph F. and Nancy P. Keithley Professor of Electrical Engineering. He is the Director of the MIT Microsystems Technology Laboratories. His research interests include low-power digital integrated circuit design, wireless microsensors, ultra-wideband radios, and emerging technologies. He is a co-author of *Low Power Digital CMOS Design* (Kluwer Academic Publishers, 1995), *Digital Integrated Circuits* (Pearson Prentice-Hall, 2003, 2nd edition), and *Sub-threshold Design for Ultra-Low Power Systems* (Springer 2006). He is also a co-editor of *Low Power CMOS Design* (IEEE Press, 1998), *Design of*

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.

High-Performance Microprocessor Circuits (IEEE Press, 2000), and *Leakage in Nanometer CMOS Technologies* (Springer, 2005).

Dr. Chandrakasan has been a co-recipient of several awards including the 1993 IEEE Communications Society's Best Tutorial Paper Award, the IEEE Electron Devices Society's 1997 Paul Rappaport Award for the Best Paper in an EDS publication during 1997, the 1999 DAC Design Contest Award, the 2004 DAC/ISSCC Student Design Contest Award, the 2007 ISSCC Beatrice Winner Award for Editorial Excellence and the 2007 ISSCC Jack Kilby Award for Outstanding Student Paper. He has served as a technical program co-chair for the 1997 International Symposium on Low Power Electronics and Design (ISLPED), VLSI Design '98, and the 1998 IEEE Workshop on Signal Processing Systems. He was the Signal Processing Sub-committee Chair for ISSCC 1999–2001, the Program Vice-Chair for ISSCC 2002, the Program Chair for ISSCC 2003, and the Technology Directions Sub-committee Chair for ISSCC 2004–2008. He was an Associate Editor for the IEEE JOURNAL OF SOLID-STATE CIRCUITS from 1998 to 2001. He served on SSCS AdCom from 2000 to 2007 and he was the meetings committee chair from 2004 to 2007. He is the Technology Directions Chair for ISSCC 2009.

Authorized licensed use limited to: MIT Libraries. Downloaded on January 17, 2009 at 11:25 from IEEE Xplore. Restrictions apply.