0.23 V Sample-Boost-Latch Based Offset Tolerant Sense Amplifier

Dhruv Patel, *Student Member, IEEE*, Manoj Sachdev, *Fellow, IEEE*

Abstract—An offset tolerant SRAM Sense Amplifier (SA) deployed with Sample-Boost-Latch (SBL) technique to facilitate both common mode and differential mode boosting is proposed. The common mode boosting enables proposed SA, SBLSA to operate deeper into the subthreshold regime whereas the differential mode boosting helps tolerate undesirable mismatch conditions by driving cross-coupled inverters with boosted differential voltages. The proposed boosting circuit first samples the bitline voltages and then isolates the sampled voltages from the highly capacitive bitlines followed by the boosting phase on relatively much less internal input capacitances of the latch. Arrays of SBLSA and conventional voltage latch SA (VLSA) were fabricated in 65nm-GP CMOS (512 each). At 25 ˚C, SBLSA achieves 23.3% reduction in directly measured standard deviation of input-referred offset (across 16 ICs, 8192 SAs) and 38.5% improvement (typical IC) in sensing delay at 0.3 V. SBLSA offers reliable operation across entire temperature range of −25 ˚C to 85 ˚C at 0.35 V whereas it is 0.41 V for VLSA. Finally, SBLSA operates reliably at the minimum supply of 0.23 V at 25 ˚C which is 30 mV less compared to VLSA.

Index Terms— Offset Tolerant Circuits, SRAM Sense Amplifier, SRAM Yield Improvement, Subthreshold Circuits, Variation Tolerant Circuits

I. INTRODUCTION

Low-power and reliable SRAM is in high demand for SoCs used in battery operated circuits, bio-implantable and IoT devices. This work is specifically targeted towards improving efficiency, speed and reliability of SRAMs in such SoCs operating at sub-threshold voltages and wide range of environment conditions (i.e. temperature). Overall SRAM performance is heavily limited by the input-referred offset of the sense amplifier (SA) as it requires minimum worst-case differential voltage developed on highly capacitive bitlines greater than the SA's input-referred offset [1]. Out of the two popular conventional SA schemes, Current Latch SA (CLSA) and Voltage Latch SA (VLSA), VLSA offers ~3x tighter offset distribution compared to prior one within same area budget making it a good benchmark scheme [2]. Recent SA works making further improvements in offset tolerance while benchmarking VLSA typically use differential pre-amplifier [3, 4], threshold matching by capacitive storage [5], and deploying re-configurable redundancy [6] at the cost of routing complexity, increased area and/or imposing built-in-self-test for fuse-based best configuration selection. More importantly, they do not offer sufficient offset tolerance and reliability across wider temperature range in sub-threshold regime, and hence weak candidates for the applications targeted in this work. To increase offset tolerance and sensing speed in subthreshold regime while achieving wider reliability coverage across temperature, this work proposes Sample-Latch-Boost based Sense Amplifier (SBLSA) scheme.

II. SBLSA WITH SAMPLE-BOOST-LATCH TECHNIQUE

Unlike previous works of [3] and [4], proposed SA scheme is capable of both differential-mode boosting (DMB) and common-mode boosting (CMB) of bitline large-signal locally on less capacitive internal input nodes of the proposed SBLSA (Fig. 1) at a reasonable sensing area penalty of 12% compared to VLSA (Fig. 2). This work modifies the VLSA by adding single-shot charge pump based boosting circuit comprised of 0.84 fF MOS capacitor (adding 48% more capacitance on Q/QB) and 2 near-minimum sized switches on both Q/QB nodes. Additionally, this work modifies the cross-coupled inverter stage by judicially using bitline signals as sources for P1/P2 to further tolerate offset by gaining higher differential current in regeneration branches. In a full SRAM system, bitline capacitance is up to few hundreds of fF and therefore, can safely be assumed as an ideal supply for SBLSA.
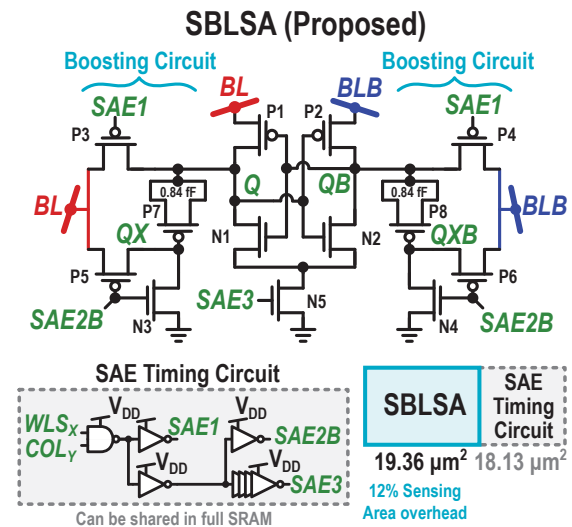

Fig. 1.   Schematic of proposed SBLSA deployed with boosting circuitry.
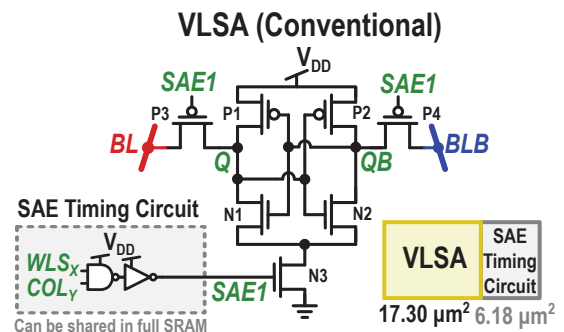

Fig. 2.   Schematic of conventional Voltage Latch SA (VLSA).

Engineering at the University of Toronto, Toronto, ON, M5S 3G4, Canada (e-mail: dhruv.patel@isl.utoronto.ca).
Manoj Sachdev is with the Department of Electrical and Computer Engineering at the University of Waterloo, Waterloo, ON, N2L 3G1, Canada. (e-mail: msachdev@uwaterloo.ca).

Dhruv Patel was with the University of Waterloo, Waterloo, ON, N2L 3G1, Canada. He is now with the Department of Electrical and Computer
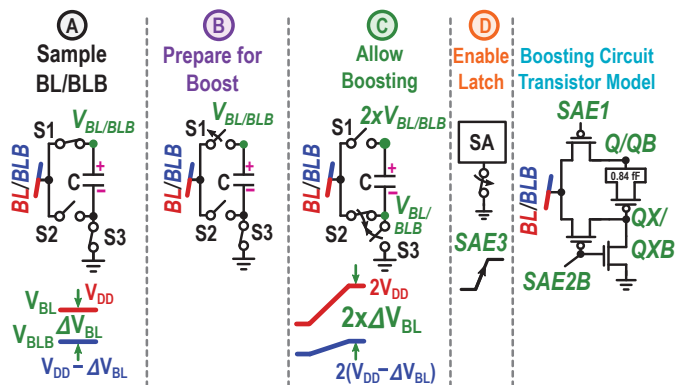
Fig. 3.   Boosting circuit concept comprised of four phases (A-D) and its transistorized circuit model used in SBLSA sensing scheme.

The boosting circuit in SBLSA has four phases (A-D) which are conceptualized in Fig. 3 with its transistorized circuit model. The functionality of SBLSA across each phase can be explained as following: **A) Sample BL/BLB:** Initially with SAE1 and SAE3 being low (P3/P4 ON, N5 OFF) and SAE2B being high (P5/P6 OFF, N3/N4 ON), allow BL/BLB to charge Q/QB and MOS capacitor devices (P7/P8) to $V_{BL/BLB}$. **B) Prepare for Boost:** Isolate BL/BLB from Q/QB by making SAE1 high (P3/P4 OFF). **C) Allow Boosting:** By this moment, MOS capacitors have potential of $V_{BL/BLB}$ on Q/QB while QX/QXB nodes are discharged to GND. Making SAE2B low (P5/P6 ON, N3/N4 OFF), allows BL/BL to charge up QX/QXB nodes to $V_{BL/BLB}$. Since the potential across capacitors can't change instantaneously, the un-driven Q/QB nodes boost up in accordance with QX/QXB. Theoretically, making QX/QXB rise to $V_{BL/BLB}$ allows Q/QB to boost up to $2xV_{BL/BLB}$ giving 2xCMB. Assuming $V_{BL} = V_{DD}$ and $V_{BLB} = V_{DD} - \Delta V_{BL}$, DMB also results in $2x\Delta V_{BL}$. **D) Enable Latch & Resolve:** Finally, SAE3 is asserted high to start the regeneration process and resolve output. Due to the addition of the boosting circuit, unlike in VLSA, cross-coupled inverter gates in SBLSA are driven with higher gate-to-source voltage owing to CMB and higher differential overdrive voltage owing to DMB. This makes SBLSA capable of operating deeper into the subthreshold region with increased offset tolerance.
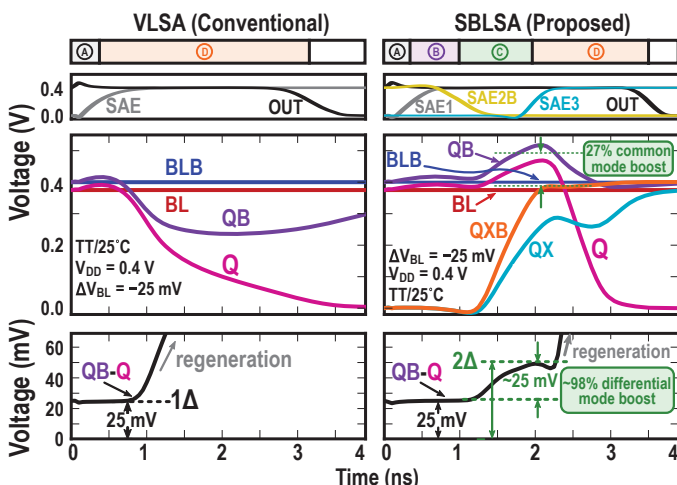

Fig. 4.   Post-extracted transient simulation comparison between VLSA and SBLSA at V_{DD} of 0.4 V with applied ΔV_{BL} of -25 mV.

Boosting scheme, however, impose non-ideality due to charge sharing between MOS capacitors and internal capacitances of the latching element. Moreover, as CMB raise

Q/QB nodes towards $V_{DD} + V_{TH-P3/P4}$, it can partially turn on P3/P4 devices leaking boosted charge back into bitlines and also slightly forward bias the *pn* junctions at the p+ diffusions on the Q/QB nodes leaking boosted charge back to V_{DD} through *n-well*; all resulting in reduced CMB and DMB. The MOS capacitor size was chosen as a compromise between boosting benefits and SA area penalty. The SAE phase sequence (A-D) is simply generated by few inverters. The transient operation (post-extracted) of the SBLSA indicating both CMB and DMB of 27% and 98% is shown in Fig. 4 while making a comparison with VLSA, respectively. The simulations were performed at 0.4 V and 25 ˚C/TT corner with ΔV_{BL} of -25mV.
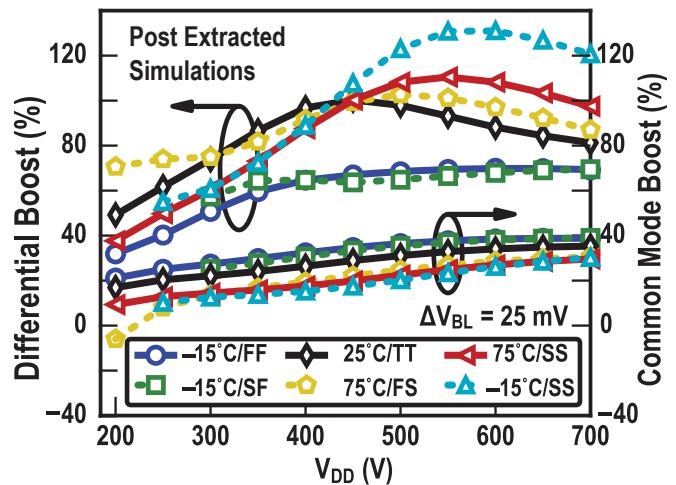

Fig. 5.   Post-extracted characterization: Differential and Common mode large-signal boosting for SBLSA across V_{DD} and worst-case corners.

Fig. 5 shows the post-extracted simulations characterizing DMB and CMB across V_{DD} and worst-case corners validating functionality of the boosting circuit. For 25 ˚C/TT corner at $\Delta V_{BL} = 25$ mV, DMB is between $50\% - 98\%$ and CMB is between $18\% - 25\%$ for 0.2 V $-$ 0.4 V supply voltage range, respectively. The average total power consumption of SBLSA (internal circuit + output loads + loading at switches in boosting circuity) is increased by $2x - 3x$ compared to VLSA under the assumption of 100% data activity as shown in Fig. 6. However, the Energy-Delay-Product is lower for SBLSA than VLSA for 0.3 V $-$ 0.6 V supply voltage range (owing to DMB peaking) and comparable elsewhere (owing to reduced sensing delay improvement due to diminishing DMB) with the same activity rate assumption as before. Moreover, for the practical read activity of $10\% - 20\%$, power consumption of SBLSA in SRAM would be much lower compared to SRAM's leakage power in the retention mode.
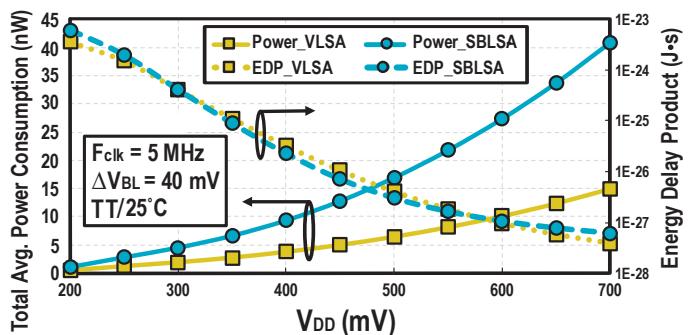

Fig. 6.   Total average power consumption and Energy-Delay-Product comparison between VLSA and SBLSA (post extracted simulations).

III. MEASUREMENTS

Shmoo plots of Frequency-V_{DD} (with 0.5 MHz and 10 mV step size at 25 °C) and Temperature-V_{DD} (with 10 °C and 10 mV step size at 5 MHz) were measured comparing both schemes on a typical die at ΔV_{BL} of ±40 mV with the passing threshold of <0.8% SA read error rate. As depicted in Fig. 7, at a given V_{DD}, SBLSA is capable of operating at higher frequency and wider temperature range compared to VLSA. Temperature-V_{DD} shmoo plot indicates that SBLSA operates reliably across entire temperature range of −25 °C to 85 °C at 0.35 V whereas it is 0.41 V for VLSA. At 25 °C, SBLSA can operate reliably at minimum V_{DD} of 0.23 V which is 30 mV lower compared to VLSA. Sensing delays from maximum possible frequency of operation at a given V_{DD} were extracted (Sensing Delay (ns) = $1/(2 \cdot \text{Freq}_{max})$) from Frequency-$V_{DD}$ shmoo plot for both sensing schemes where the relative trends closely match with the post-extracted simulations as shown in Fig. 8. SBLSA achieves 38.5% improvement in sensing delay at 0.3 V.
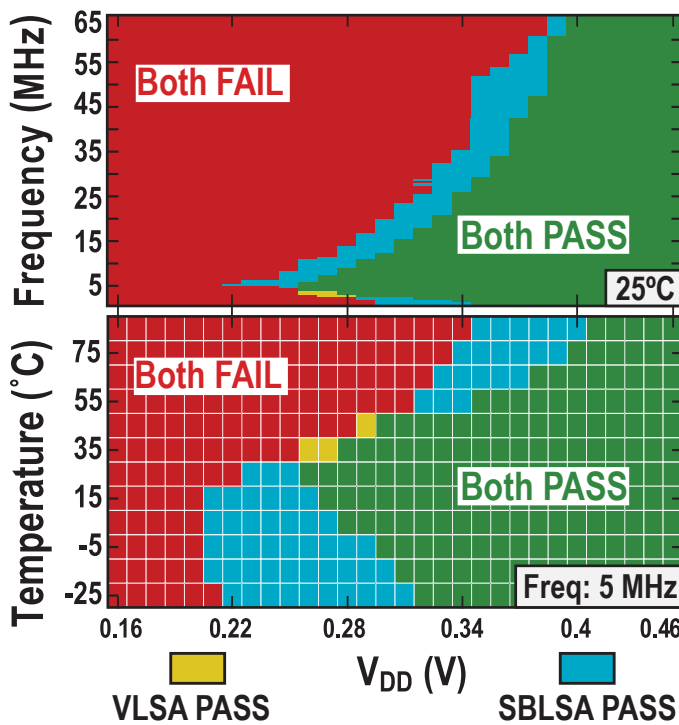
Cumulative distribution function (CDF) of yield curves were measured across V_{DD} (0.23 V to 0.7 V) and temperature (−25 °C to 85 °C) at 5 MHz while sweeping ΔV_{BL} in [0, +5mV, −5mV, … +50mV, −50mV] pattern. Yields at given ΔV_{BL} were combined from 16 dice (8192 SAs each) giving CDF plot at a given V_{DD} and temperature. The standard deviation of the input-referred offsets (Std_{OS}) were extracted from the probability density curves derived from their respective CDF curves; example shown in Fig. 9. Std_{OS} vs V_{DD} curves in Fig. 10 show that SBLSA has lower Std_{OS} compare to VLSA across all V_{DD} at respective temperatures. Complying with DMB simulations, Std_{OS} of SBLSA is relatively lower where the DMB is peaking (0.4 V−0.5 V) with sufficient stability across temperature. Both schemes had their Std_{OS} decreased from low-to-high temperatures at a given V_{DD}. At 25 °C, SBLSA had 23.3% and 24.7% (peak) improvement at 0.3 V and 0.4 V in Std_{OS} as shown in Fig. 11 (a), respectively. According to the work of Pileggi et. al. [7], since VLSA's offset is mainly determined by its NMOS pair accurately following Pelgrom's mismatch model, the maximum possible offset improvement with 12% (incurred sensing area penalty in proposed SBLSA) additional area in NMOS pair of VLSA is only 5.5%, whereas SBLSA takes this improvement to ~4x higher. The Inter-Die standard deviation of Std_{OS} (ID-$\text{Std}_{\text{Std-OS}}$) was also extracted over 512 SAs on each of the 16 dice. From Fig. 11 (b), ID-$\text{Std}_{\text{Std-OS}}$ is within 0.25 mV−1.25 mV for SBLSA (lower or comparable to VLSA) under all testing conditions further validating inter-die consistency of offset tolerance with proposed SBLSA scheme.

Fig. 7.   Measured Frequency-V_{DD} (top) & Temperature-V_{DD} (bottom) shmoo plots on a typical die for VLSA and SBLSA. Measurements performed at constant ΔV_{BL} of ±40 mV where the worst-case yield was considered.
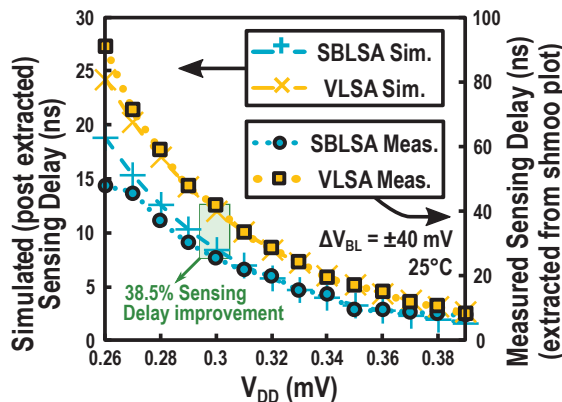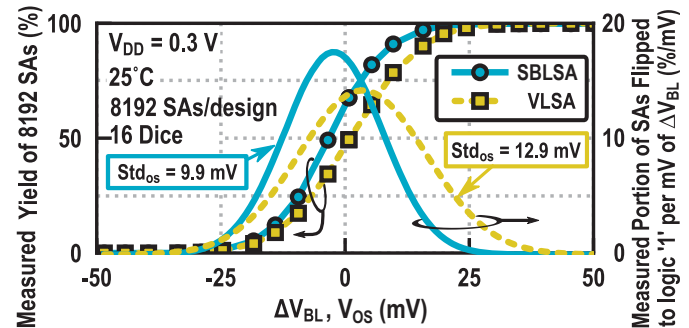

Fig. 8.   Measured (on a typical die) and post-extracted simulation sensing delay vs V_{DD} comparison between VLSA and SBLSA.

Fig. 9.   Measured cumulative distribution and respective probability density function of input-referred offset statistics of VLSA and SBLSA across 16 dice (8192 SAs). Measured at V_{DD} = 0.3 V, 25 °C and 5 MHz of clock frequency.
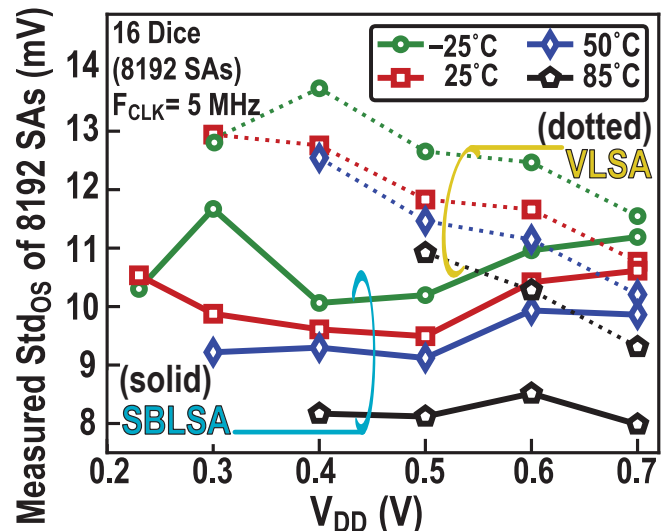

Fig. 10.   Measured Std_{OS} of VLSA and SBLSA across 16 dice (8192 SAs).

Similar to [2], addressable arrays (512 each) of both VLSA and SBLSA (layout shown in Fig. 12) were prototyped in 65nm-GP CMOS for direct input-referred offset characterization with BL/BLB driven and swept externally. Fig. 13 shows the die photograph annotated with top level test-chip blocks. The SAE timing circuit was also arrayed along with its respective sensing scheme to include the impact of timing variation; especially to add additional rigor for boosting circuit. However, in full SRAM systems, all three SAE timing signals can be shared across SBLSAs similar to how a single SAE signal would be shared across VLSAs. The timing sequence of SBLSA in full SRAM is shown in Fig. 14. It shows that SAE1 to SAE3 delay can be absorbed while ΔV_{BL} is developed on bitlines. Note that in phase C, as SAE2B makes 1→0 transition with ΔV_{BL} still under development, BL/BLB now charge up QX/QXB nodes. In parallel, the Q/QB nodes gets boosted in accordance with QX/QXB taking full benefit of discharged bitlines during entire WL pulse. Finally, the comparison of SBLSA with the state-of-the-art SA schemes is shown in Fig. 15.
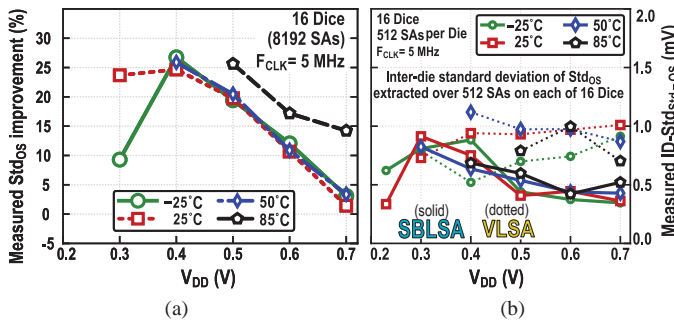

Fig. 11. (a) Std_{OS} improvement from Fig. 10 across V_{DD} and Temperature. (b) Measured ID-Std$_{Std-OS}$ extracted over 512 SAs on each of 16 dice.
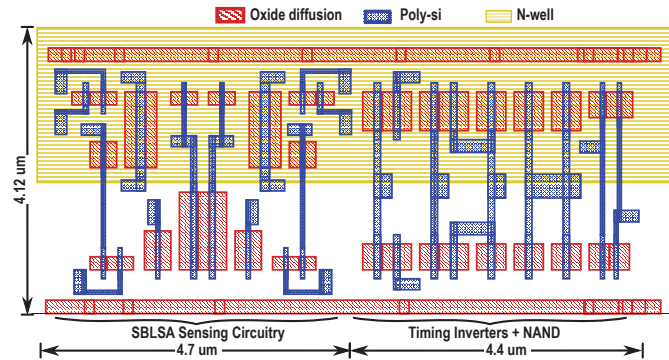

Fig. 12. Layout of proposed SBLSA with SAE timing circuitry.

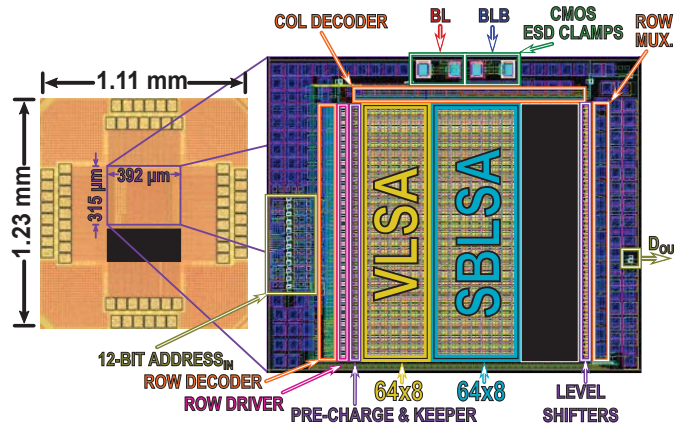
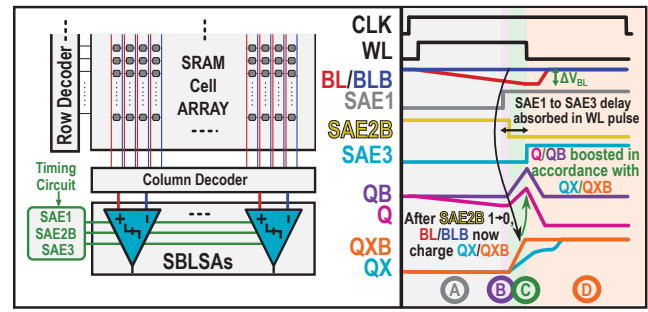Fig. 13. 65nm CMOS test-chip with 512 arrays of VLSA & SBLSA.

Fig. 14. Timing sequence concept of proposed SBLSA in full SRAM.

IV. CONCLUSION

The Sample-Boost-Latch based SRAM SA, SBLSA was proposed. It achieves 23.3% and 38.5% measured improvement in standard deviation of input-referred offset and sensing delay in 65nm-GP CMOS compared to conventional VLSA, respectively. It is also capable of operating at 30 mV lower supply (at 0.23 V) at 25 ˚C and offers wider reliable operation coverage across temperature compared to VLSA.

SA comparison	[3] ISSCC'14	[4] A-SSCC'16	[5] JSSC'16	[6] Sym. VLSI'15	[8] JSSC'14	[SBLSA] This Work
Tech. Node	28nm	28nm HPM	28nm HP	28nm FDSOI	65nm LP	65nm GP
Sensing Scheme	Small sig. pre-amp, Auto-zero offset cancellation	Small sig. pre-amp, self-timed	MOS cap based threshold matching	Reconfig. redundancy w/ BIST config.	CLSA w/ Body-Bias offset calibration	Large sig. diff. & comn. mode boost, bitlines as SA supply
# of Devices	10T + 2 MOM caps	15T	11T+ 2 MOS caps + 5 INV	24 T + 1 fuse	15T+2 NOR + 2 NAND +3 INV +1 Latch	11T + 2 MOS caps
SA Area	1x [a.] + 2 MOM caps [b.]	1x [a.]	3.2% [a., d.]	1x [a.,] + fuse	3.5% [c., d,]	1.12x [a.]
Offset reduction	N/A	22% [a.] @0.45 V, 25°C	49% [a.] @0.5 V, 85°C	~2x [a.] @0.4 V, 25°C	50% [c.] @ 1.2 V	23.3% in Std. of offset @0.3 V, 25°C
Sensing Delay reduction [a.]	34% @1 V, 27°C	13% @0.6V	N/A	N/A	N/A	38.5% @0.3V, 25°C
V_{DD-min} @25°C	N/A	450 mV	500 mV	400 mV	370 mV	230 mV
V_{DD-min} w/ Temp. range	−5°C to 85°C @ 1V, 1.8GHZ	N/A	N/A	−25°C to 85°C @0.4 V	N/A	−25°C to 85°C @0.35 V, 5MHz
SA Samples	352 (22 dice)	432 (6 dice)	2k (1 die)	19k (19 dice)	512 (1 IC)	8k (16 dice)
Char. Type	Full SRAM	Full SRAM	Full SRAM	SA Array	Full SRAM	SA Array

[a.] w.r.t VLSA implemented in respective work
[b.] Penalty of 2 MOM caps in 2 metal layers only (not in active region)
[c.] w.r.t CLSA implemented in respective work
[d.] Overall area penalty for full 128 kB SRAM

Fig. 15. Comparison with the state-of-the-art offset tolerant SRAM SAs.

REFERENCES

[1] B. Wicht, *et al.,* "Yield and speed optimization of a latch-type voltage sense amplifier," in *IEEE JSSC*, vol. 39, no. 7, pp. 1148-1158, July 2004.

[2] M.H. Abu-Rahma, *et al.,* "Characterization of SRAM sense amplifier input offset for yield prediction in 28nm CMOS," *2011 IEEE CICC*, 2011, pp. 1-4.

[3] B. Giridhar, *et al.,* "13.7 A reconfigurable sense amplifier with auto-zero calibration and pre-amplification in 28nm CMOS," *2014 IEEE ISSCC*, 2014, pp. 242-243.

[4] P. F. Chiu, *et al.*, "A double-tail sense amplifier for low-voltage SRAM in 28nm technology," *2016 IEEE A-SSCC*, 2016, pp. 181-184.

[5] M. E. Sinangil, *et al.*, "A 28 nm 2 Mbit 6 T SRAM With Highly Configurable Low-Voltage Write-Ability Assist Implementation and Capacitor-Based Sense-Amplifier Input Offset Compensation," in *IEEE JSSC*, vol. 51, no. 2, pp. 557-567, Feb. 2016.

[6] Mahmood Khayatzadeh, *et al.* "A reconfigurable sense amplifier with 3X offset reduction in 28nm FDSOI CMOS," *2015 Symposium on VLSI Circuits (VLSI Circuits)*, Kyoto, 2015, pp. C270-C271.

[7] L. Pileggi, *et al.*, "Mismatch analysis and statistical design at 65 nm and below," *2008 IEEE CICC*, San Jose, CA, 2008, pp. 9-12.

[8] Y. Sinangil and A. P. Chandrakasan, "A 128 Kbit SRAM With an Embedded Energy Monitoring Circuit and Sense-Amplifier Offset Compensation Using Body Biasing," in *IEEE JSSC*, vol. 49, no. 11, pp. 2730-2739, Nov. 2014.