An Area-Efficient Low-Offset Sense Amplifier for In-Memory Computation

A REPORT

Submitted by

Mosom Jana (EE17B110)

in partial fulfillment of the requirements

for the award of the degree

DUAL DEGREE in Electrical Engineering

Under the guidance of

Dr. Janakiraman Viraraghavan

DEPARTMENT OF ELECTRICAL ENGINEERING

INDIAN INSTITUTE OF TECHNOLOGY MADRAS

CHENNAI-600036

JUNE 2022

CERTIFICATE

This is to certify that the report entitled "**An Area-Efficient Low-Offset Sense Amplifier for In-Memory Computation**" submitted by **Mosom Jana** to the **Indian Institute of Technology, Madras,** for the award of the degree of **Bachelor of Technology (B.Tech.)** and **Master of Technology (M.Tech.)** in **Electrical Engineering** is a bonafide record of academic work carried out by him under my supervision. The contents of the work, in full or in parts, are not submitted to any other institute for the award of any degree or diploma.

Prof. Janakiraman Viraraghavan
Project Guide and Assistant Professor
Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai – 600 036.

Prof. R David Koilpillai
Professor and Head
Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai – 600 036.

Place: Chennai

Date: June 2022

ACKNOWLEDGEMENTS

It gives me a sense of fulfillment to write a few words to acknowledge people without whose support this report would not have been possible. I heartfully express my gratitude to my guide **Dr. Janakiraman Viraraghavan**, Assistant Professor, Integrated Circuits and Systems group, Department of Electrical Engineering, IIT Madras, for his valuable guidance. I would also like to express my gratitude to **Ashwin Balagopal,** Research Scholar under Dr. Janakiraman Viraraghavan for taking the time to help me with my work, particularly with the software.

I would like to acknowledge **Dr. David Koilpillai**, Professor and Head of Department, Electrical Engineering, IIT Madras, for allowing me to undertake and extend this project.

I am thankful to all my friends who have been a source of support and help whenever needed. I am grateful to my parents, who have always encouraged me.

ABSTRACT

Keywords: Autozeroing Sense Amplifier, Flash ADC, Offset Behavior.

The present IMC chip under our iCS group has a Sense Amplifier (comparator) that has been observed to behave peculiarly, both in hardware and in simulations, particularly in its offset behavior. An attempt is made to address this abnormality by understanding the present design and trying to improve the design to meet our targets, which are primarily offset and area constrained. The limitation of improving the existing design is highlighted in terms of the offset and area trade-offs. A simplistic yet novel Sense Amplifier is presented which uses an offset minimization technique called Autozeroing. This enables the new design to populate lesser On-Chip space while delivering far lower offsets than the improved version of the existing Sense Amplifier. The layout for the same is explored with post-layout simulations validating the robustness of the new Sense Amplifier across corners, variations, and noise turbulences. The 15-Sense Amplifier array for a 4-bit Flash ADC is laid out and its performance is evaluated.

TABLE OF CONTENTS

ABBREVIATIONS

MC	Monte Carlo
SA	Sense Amplifier
ICM	Input Common Mode
ADC	Analog-to-Digital Converter
TG	Transmission Gate
CI	Charge Injection
DL	Data Line
PEX	Parasitic Extraction
AC	Alternating Current
IMC	In-Memory Computation
MAC	Multiply-and-Accumulate operation
ff, ss, tt	Fast-fast, slow-slow, typical-typical corners
MOSCAP	Metal Oxide Semiconductor Capacitor
(N)MOMCAP	(N-type) Metal-On-Metal Capacitor
N-&-PMOS(')	N-type and P-type Metal Oxide Semiconductor(s)
MBL	MAC Bitline
W, L	Width, Length
ENOB	Effective Number Of Bits

References

Deployed new SA proposed by *Mingoo Seok et al*. (2020) in "C3SRAM: An In-Memory-Computing SRAM Macro Based on Robust Capacitive Coupling Computing Mechanism", p4.

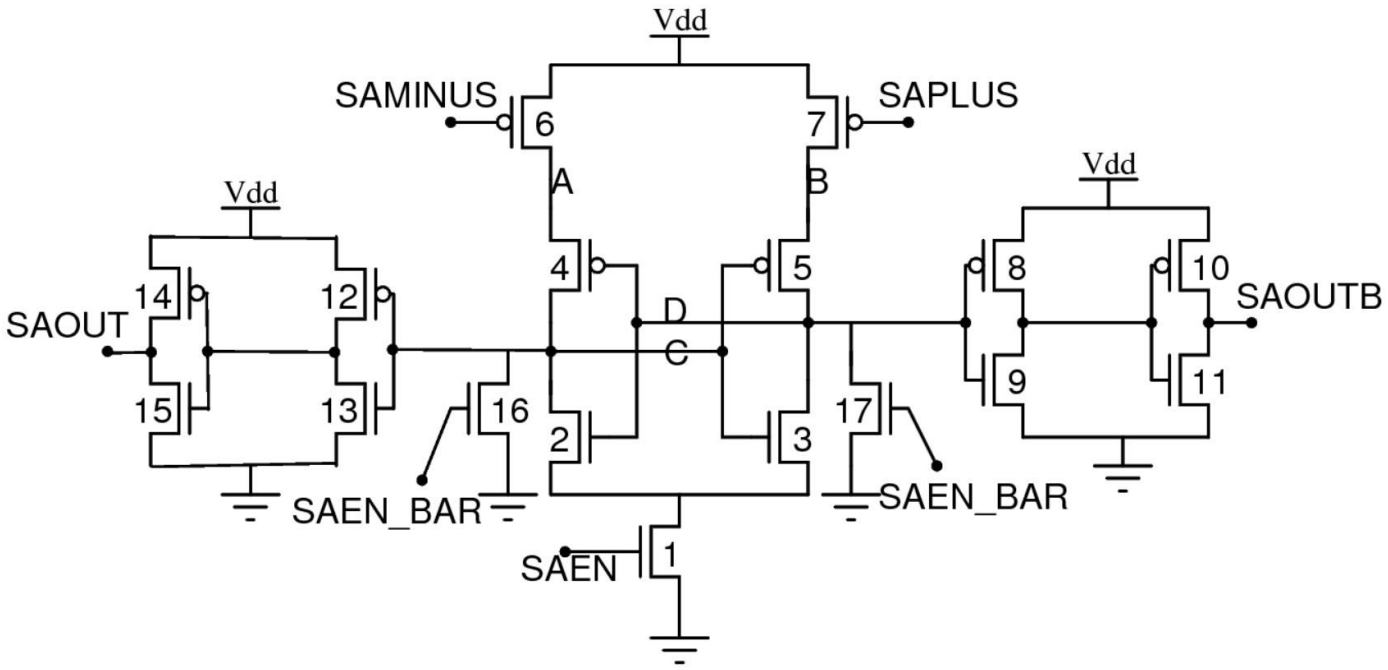
Strongarm Latch was proposed by *Behzad Razavi* (2015) in "The StrongARM Latch [A circuit for All Seasons]".

The technology used is TSMC 28nm with a VDD of 0.9V.

Software tools used: Cadence® Virtuoso®, XCircuit for drawing circuit diagrams, and Python for automating offset extraction.

Chapter 1 PREVIOUSLY DESIGNED SENSE AMPLIFIER

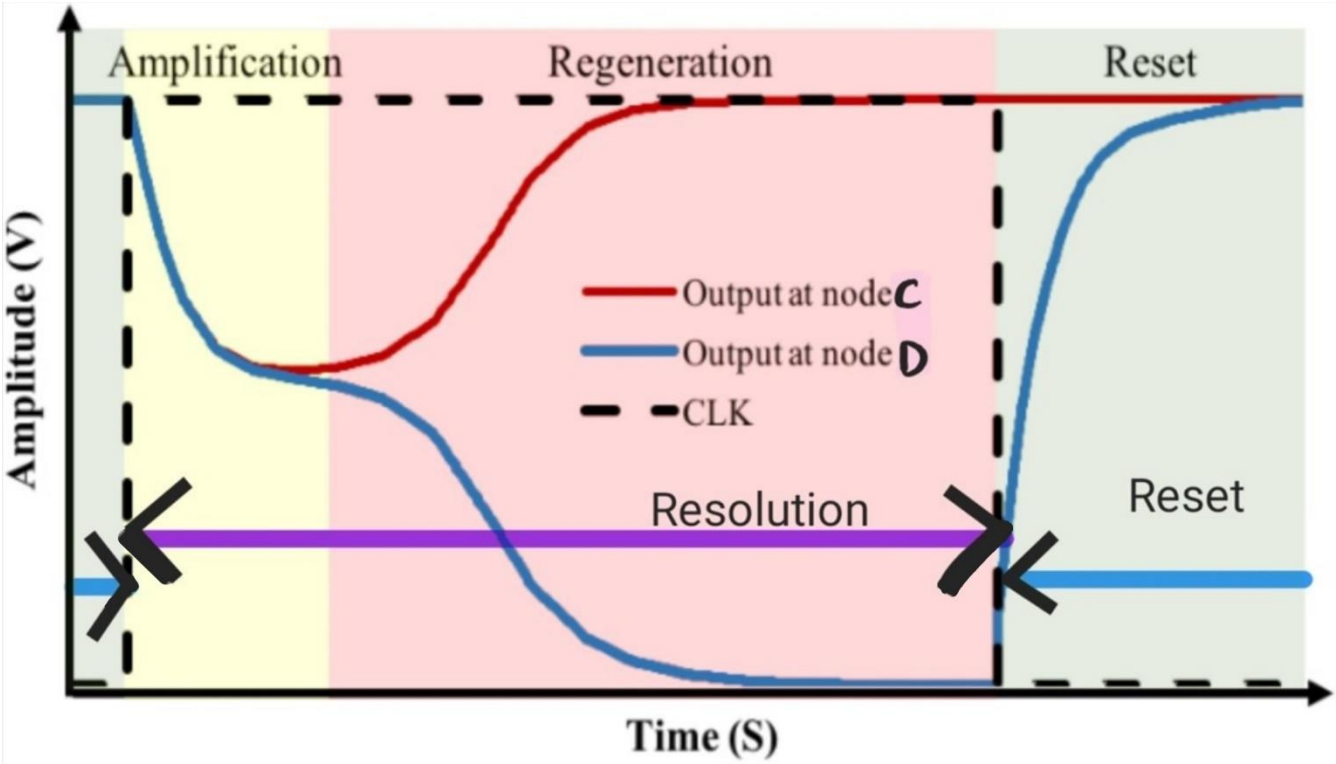
1.1 WORKING PRINCIPLE



The present Sense Amplifier (SA) that's been implemented On-Chip has the above circuit diagram. The inputs SAMINUS and SAPLUS are applied to the gates of PMOS' 6 and 7. SAEN and its inversion SAEN_ BAR (externally generated) are connected as shown, along with the output SAOUT with its inversion SAOUTB. The SA here is a modified form of the PMOS differential pair-based *Strongarm* latch. The devices (6,7), (4,5), and (2,3) are nominally supposed to be matched. This matching ensures little static offset, which is useful for our design. We describe the operation of this SA based on the Strongarm latch's usual mechanism.

The operation happens in two stages which we will call Reset (Stage 1) and Resolve (Stage 2). In the Reset phase, SAEN or SA enable is kept low. The nodes C and D are pulled low as devices 16 and 17 connecting them are strongly driven by SAEN_BAR which is at high. Devices 4 and 5 see low voltages at their gates and thereby pulling nodes A and B to low voltage levels that just turn off 4 and 5. Now, SAEN is driven high and the resolution phase starts. The nodes C and D are now floating low with 2, and 3 still off as their gates are at low voltages. Hence the path through the cross-coupled latch consisting of 2, 3, 4, and 5 is initially off during this stage. The voltages at A and B begin ramping up as the 6 and 7 inject charge on them. SAPLUS and SAMINUS are provided low enough such that 6 and 7 are ON. Let's

assume SAMINUS is less than SAPLUS. The differential rate of voltage ramping on A and B depends on the difference of magnitude between applied inputs A and B, as apparent from the schematic. This ramp up turns on 4 and 5 but at different instances since A and B are rising at unequal rates. Through our assumption, 4 turns on quicker than 5 as A ramps up faster than B, again because 6 had tuned on stronger than 7 owing to gate bias. This would mean C begins to rise before D and quicker than D since 4 turns on stronger than 5. The voltage at node C crosses the trip point of the inverter (5, 3) before that of node D crosses the trip point of the inverter (4, 2). This kicks off the positive feedback. One may visualize this regeneration by noting that as node C rises, NMOS 3 turns on stronger and pulls down node D further. This increased pull down at D turns on 4 more and C ramps up quicker. This ensures quicker ramp down at D through 3 and so forth till the nodes C and D are driven to complementary digital levels. The output inverters consisting of (14, 15), (12, 13), (8, 9), and (10, 11)) help to ensure less loading and disturbances on the cross-coupled nodes C and D.



The above plot realizes the same functionality but for a Strongarm latch based on an NMOS differential pair. Such a design is a simple device-and-signal complement of the schematic depicted above. Here, nodes C and D are reset to VDD instead of ground before the resolution phase initiates. And C and D are ramped down differentially depending on input difference instead of ramping up. The applied inputs should be high enough to turn on its NMOS differential pair.

1.2 PROBLEMS IN THE DESIGN

The On-Chip SA has a few issues that impede its performance. As we are focused on offset behavior, the offset performance was seen to vary wildly with input common mode and device mismatches, in ways that were completely unwarranted from the Strongarm latch. Before that, it helps to define the SA offset or specifically, the input-referred offset in the following way-

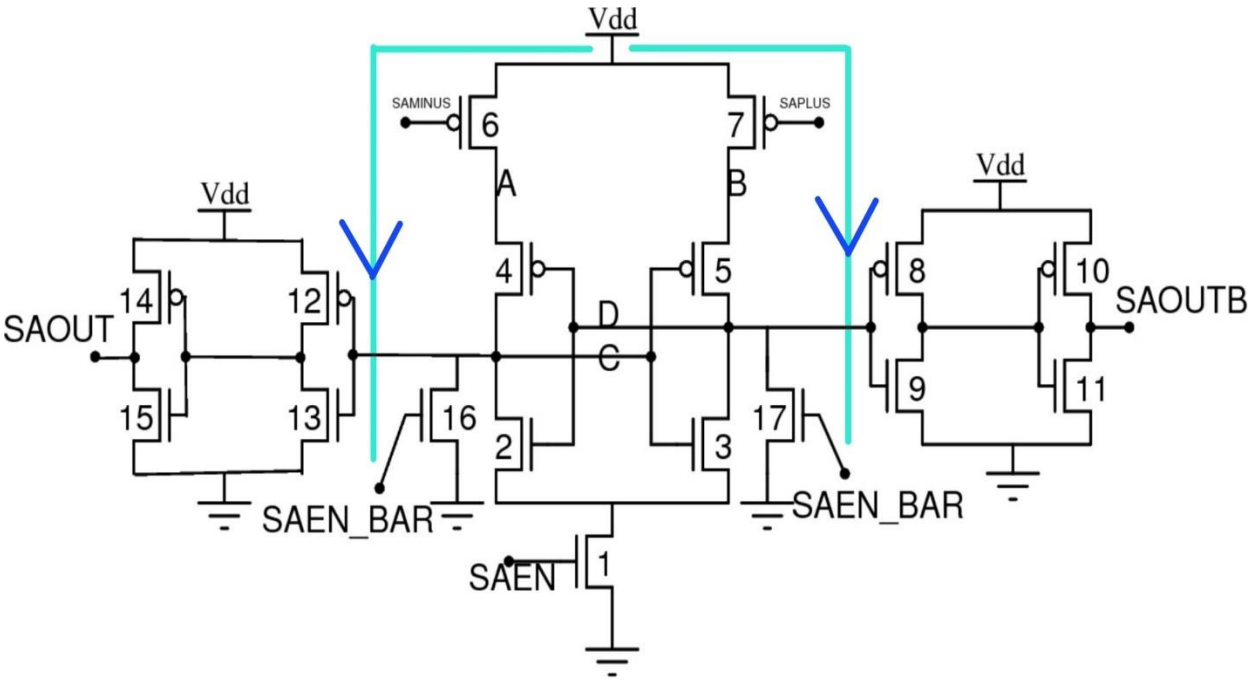
Offset is the SA's input differential (here SAPLUS – SAMINUS) at 'transition'. This 'transition' is where if the input differential is made any larger at constant input common mode, the SA outputs a high at the end of its resolve phase. Any lower, the corresponding SA output is a low.

This is the conventional offset that we already know of in comparators. But defining it this way helps follow how the offsets have been estimated on similar lines in upcoming experiments where we begin with a large positive input differential and step it down to a large negative; while inspecting the output waveform for this '*transition*' at a given input common mode level.

Now, the problems that were encountered are stated below qualitatively-

- *Inconsistency between simulation and hardware observations*: An independent MC simulation had shown that the signed value of the input-referred offset of the SA increases as the input common mode voltage is increased, whereas, for the On-Chip SA, an opposite trend was observed where the same value of offset reduces as the input common mode was increased. One might naively conclude that the 1-point sweep simulation is unreliable.
- *Variations in offset*: As the input common mode is varied, it seemed that the offset varies significantly with this variation spanning tens of millivolts. This abnormality needs to be investigated and remedied as it can severely deteriorate the performance of any SA where it creeps up.
- *Static dissipation during reset phase*: In the above schematic, during reset, there might exist a VDD to GND short circuit, depending on SAPLUS and SAMINUS that are possibly low enough to turn on devices 6 and 7. Since C and D are at low voltages during this time, 4 and 5 could turn on and with SAEN_BAR high, there is a path to ground through the PMOS' of the cross-coupled latch and into the reset NMOS'. This leads to large static dissipation and is likely to droop the VDD supply rail affecting global performance as the taped-out chip has numerous such SAs operating in parallel. This may be solved by a PMOS header switch with SAEN_BAR at its gate as it turns off the leakage path during

the duration of reset. Or, a second approach may be used where nodes C, and D are pulled-up to VDD instead of ground.



Potential VDD-GND short circuits through (6, 4, 16) and (7, 5, 17) in the SA.

1.3 ORGANIZATION OF THE REPORT

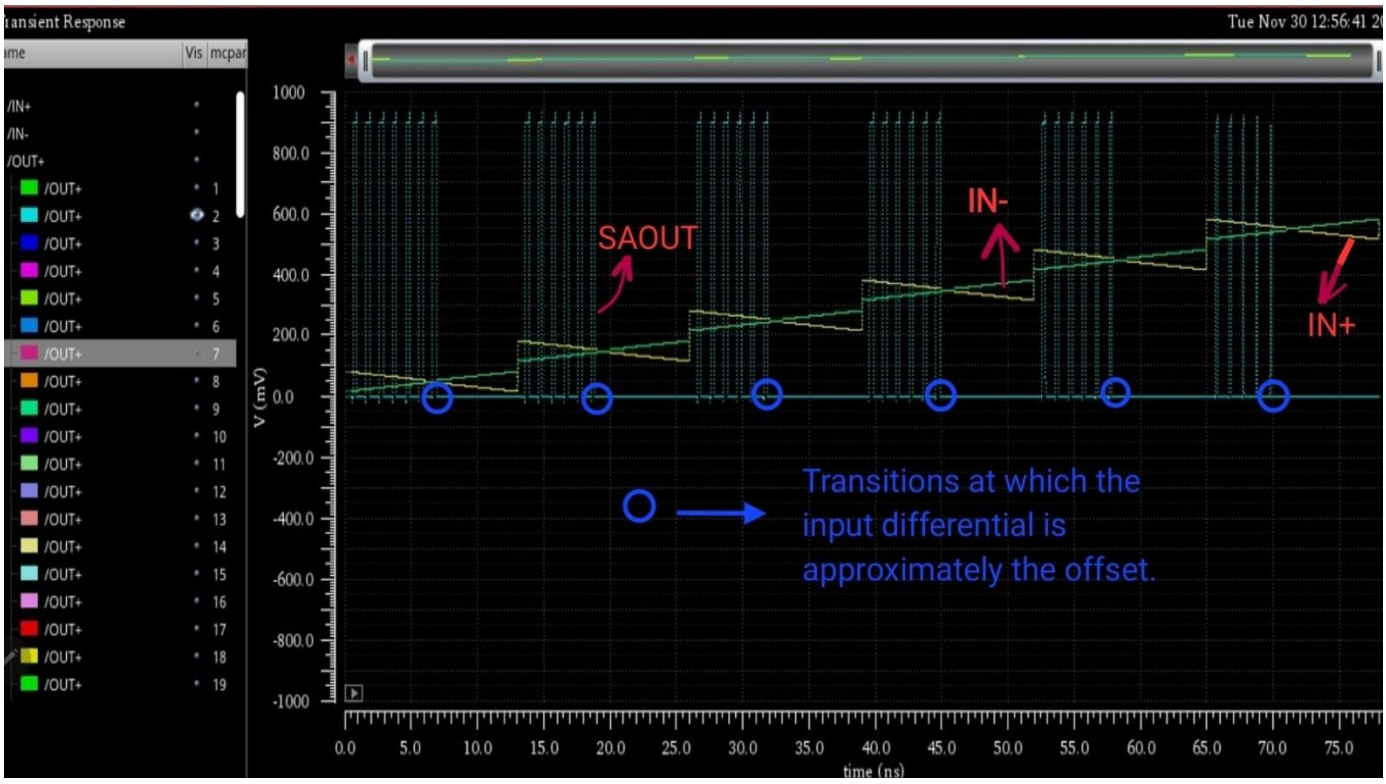
Note that the technology used here is TSMC 28nm. To document the developments made in the SA for mitigating the issues highlighted in section 1 and to implement an alternate topology for the same, the following chapters in this report have been categorized below.

- Chapter 2 explores the problems faced with the existing design and proposes the solution as a **sizing problem** that is iteratively solved over MC simulations to produce an SA that has offsets within desired bounds. We explore the layout of this SA and generate post-layout histograms to verify the same.
- Chapter 3 discusses the cons of the method adopted in Chapter 2 and uses an **auto-zeroing** structure with an **upsized inverter** in place to modestly meet our SA specifications. Recommendations are provided to help meet upgraded offset requirements for future use cases. Post-layout MC tests for a 15-SA array are done to output and validate the 15-bit thermometer code in realizing a 4-bit flash ADC.
- Chapter 4 will be a brief conclusion on this new SA, highlighting the various aspects of the design.

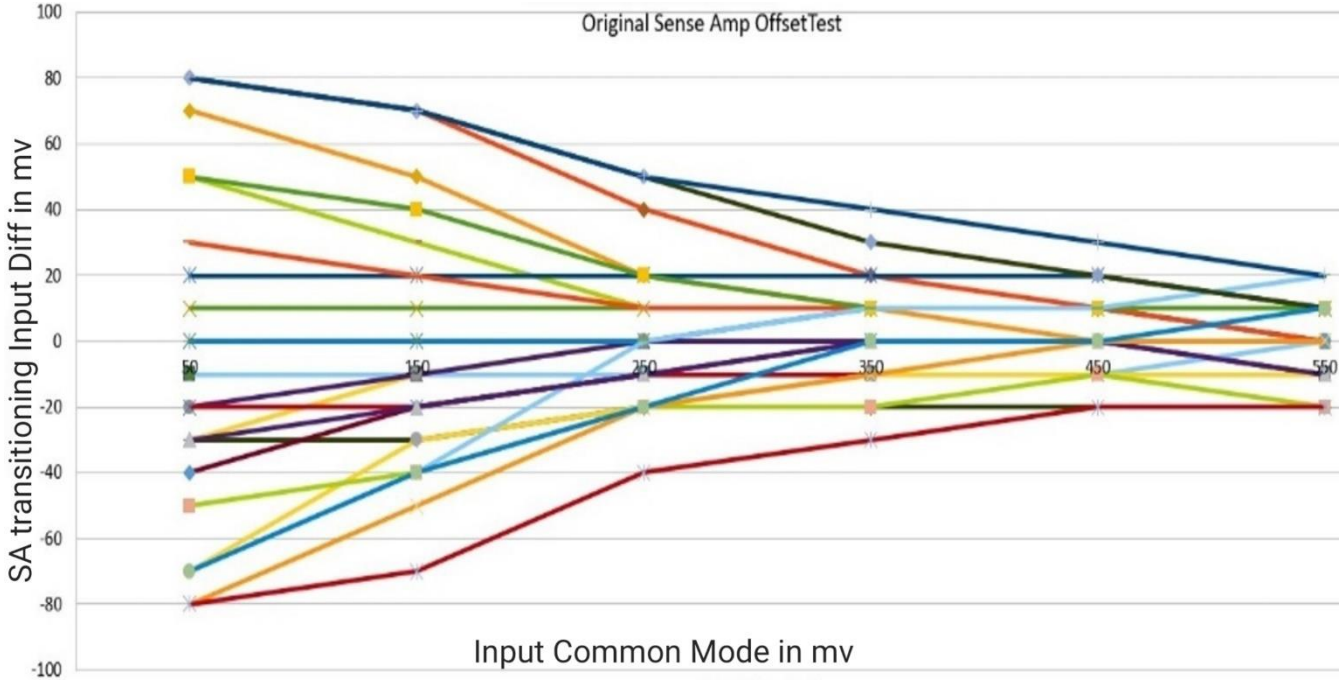
Chapter 2 SIZING THE LATCH FOR LOW OFFSET

2.1 MC TESTS FOR OFFSET ANALYSIS

To get an estimate of the offsets across variations, a simple test setup was deployed. Here we step the input common mode (ICM) from 50mv to 550mv in steps of 50mv. We restrict our input range here from 0 to 600mv roughly since if we go any higher, the input PMOS devices will barely turn on and take very long to ramp voltages at the cross-coupled nodes, thereby hampering the frequency of operation. For a fixed input common mode, the input differential is stepped from +80mv to -80mv in steps of 10mv. As described in the offset definition of section 1.2, it is inspected (and later automated) to capture at what input differential does the SA output display the *transition* of staying low as the differential is reduced thereon. Observe the plot for the SAOUT waveform. This is one of the 100 samples tested with variations at tt.

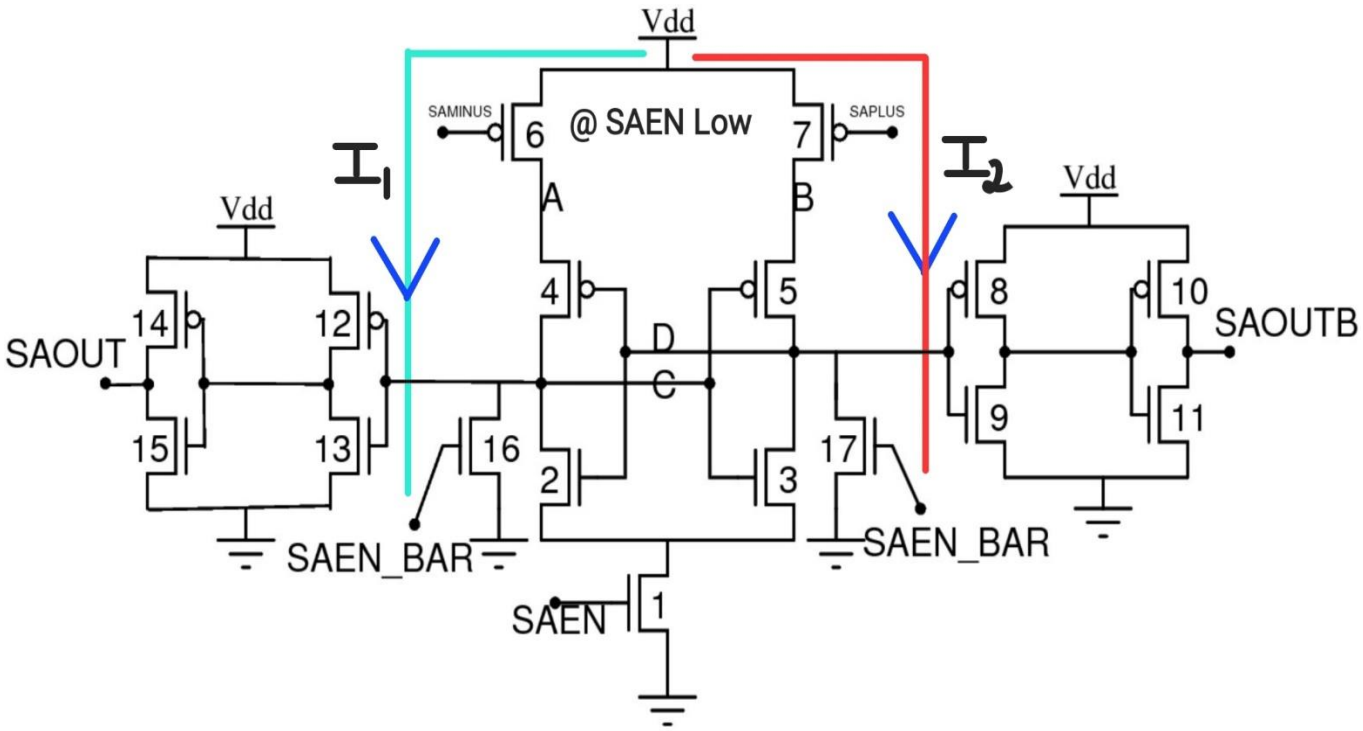


Note for the waveform IN+ is SAPLUS and IN- is SAMINUS. The offsets from 25 out of the 100 samples have been extracted by observing the transitions. For a given MC sample, the offsets are differentiated by the corresponding ICM during the transition. These 25 samples have their offsets plotted as a function of the ICM as shown below.



The offsets are rounded off to the closest higher multiple of 10mv since the differential is stepped in 10mv steps. We observe that certain samples display a wide range of offsets as the ICM is swept. Few samples have (-)80mv offset at 50mv ICM and (-)20mv at 550mv ICM. This is a bottleneck for the given SA as static offset compensation schemes are not applicable.

2.2 A REASON FOR MISBEHAVING DYNAMIC OFFSETS



As was seen in section 1.2, there are potential VDD-GND low resistance paths during the reset phase as highlighted above when SAEN goes low. These significant leakage currents I_1 need not be equal to I_2 (see fig. above) when SAPLUS doesn't equal SAMINUS since the currents through both the input PMOS' will differ nominally. This would mean the cross-coupled nodes Q and Q_bar will settle to low but somewhat unequal voltages during this reset phase owing to unequal leakage currents through them. This is primarily because both the reset NMOS' operate in deep triode and the non-zero leakage currents through them will be a strong function of their respective drain voltages which are also the voltages at the cross-coupled nodes critical to performance. The conjecture is that the misbehaving dynamic offsets have primarily got to do with the unequal starting points of Q and Q_bar at the end of reset phase and the beginning of resolution phase. We need both Q and Q_bar at 0v identically in the reset phase but a few millivolts of mismatch may be a cause of concern as is observed in this tabulation from MC.

Sample	**Dynamic offset Trend***	**Q – Q_bar in mv****
1	Increasing	6.04
2	Fixed	2.22
3	Fixed	1.85
4	Fixed	3.4
5	Fixed	1.87
6	Fixed	3.52
7	Fixed	1.92
8	Increasing	6.16
9	Fixed	2.11
10	Increasing	4.91
11	Fixed	2.22
12	Fixed	4.87
13	Fixed	4.28
14	Increasing	4.76
15	Decreasing	1.52
16	Increasing	4.66
17	Decreasing	2.55
18	Decreasing	3.43
19	Increasing	7.2
20	Fixed	3.79

*_Increasing_ is when the offset is negative at low ICMs and increases as ICM is increased by approaching 0. This was the trend seen in most MC simulations.

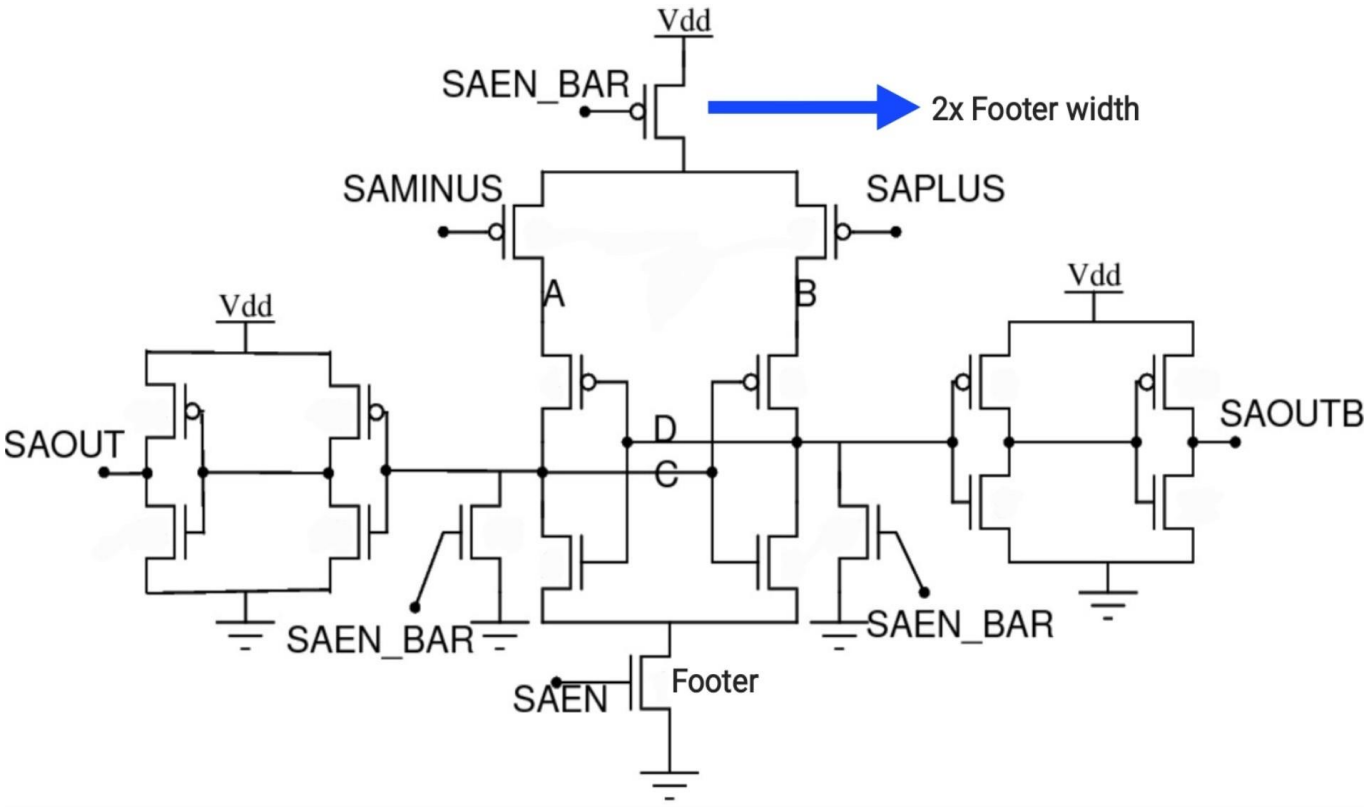
*_Fixed_ is when the offset is not changing much with ICM sweep.

*_Decreasing_ is when the offset is positive at low ICMs and decreases as ICM is increased by approaching 0. This was the trend seen in hardware for the SA.

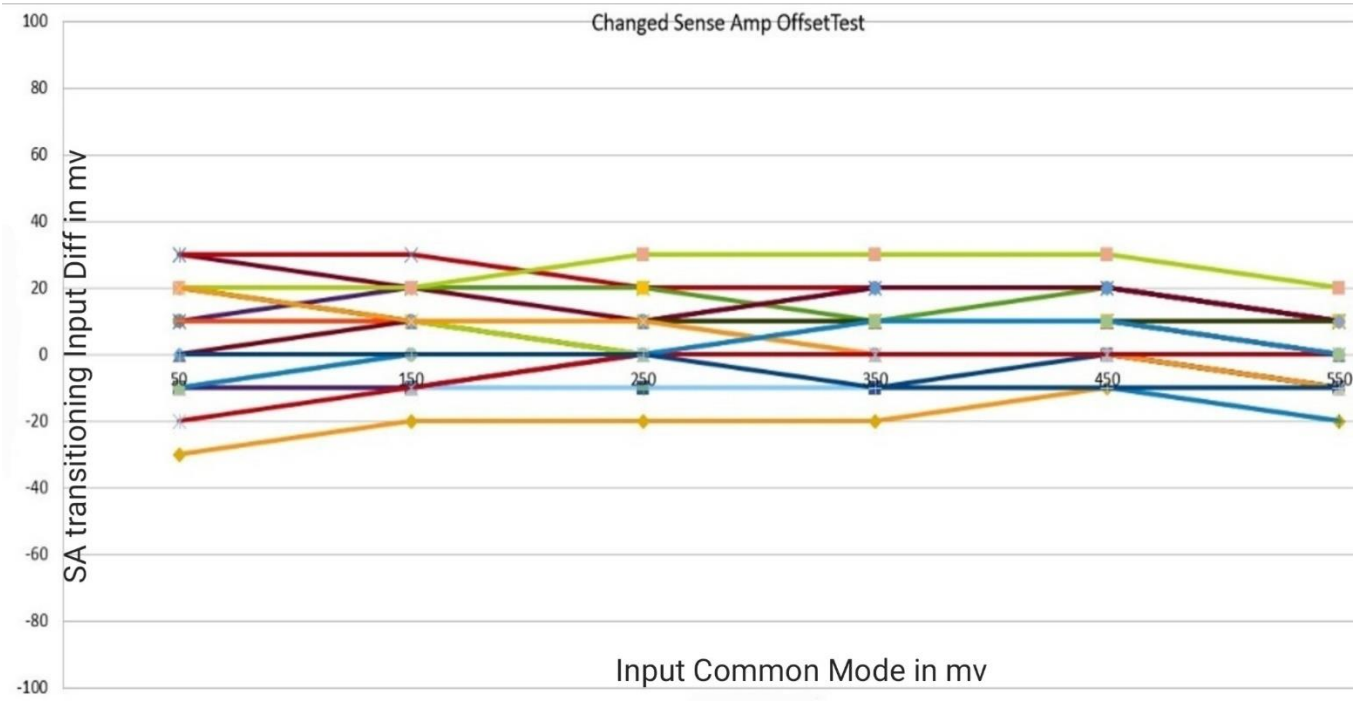
**This is cross-coupled node differential at the end of reset phase with 50mv ICM with SAPLUS at 90mv, SAMINUS at 10mv i.e., at the first step of the input differential and ICM.

The above table seems to suggest there is some correlation between the reset phase (Q Q_bar) difference and the nature of the trend in dynamic offset behavior. It appears that if the (Q Q_bar) difference is about 5mv or larger, the offsets increase with ICM. When the same difference is around 1mv or lesser as was observed for a few other samples, the offsets decrease with ICM. The (Q Q_bar) mismatch is just one of the many factors playing a role in the offset behavior apart from device-level mismatches. However, the (Q Q_bar) initial voltage difference is akin to some systematic offset induced in the SA.

To verify what happens when the (Q Q_bar) difference is made negligible by turning off the leakage path, we add a PMOS header as discussed in section 1.2 and repeat the offset tests. The circuit diagram for the same is shown below, with an additional header. In the existing design seen earlier, the PMOS width is 2um and the NMOS width is 800nm except for the footer NMOS which is 1.6um wide. The lengths of all devices are at 30nm. We retain the previous sizing to test if with the new header device, offset variations with ICM are suppressed.



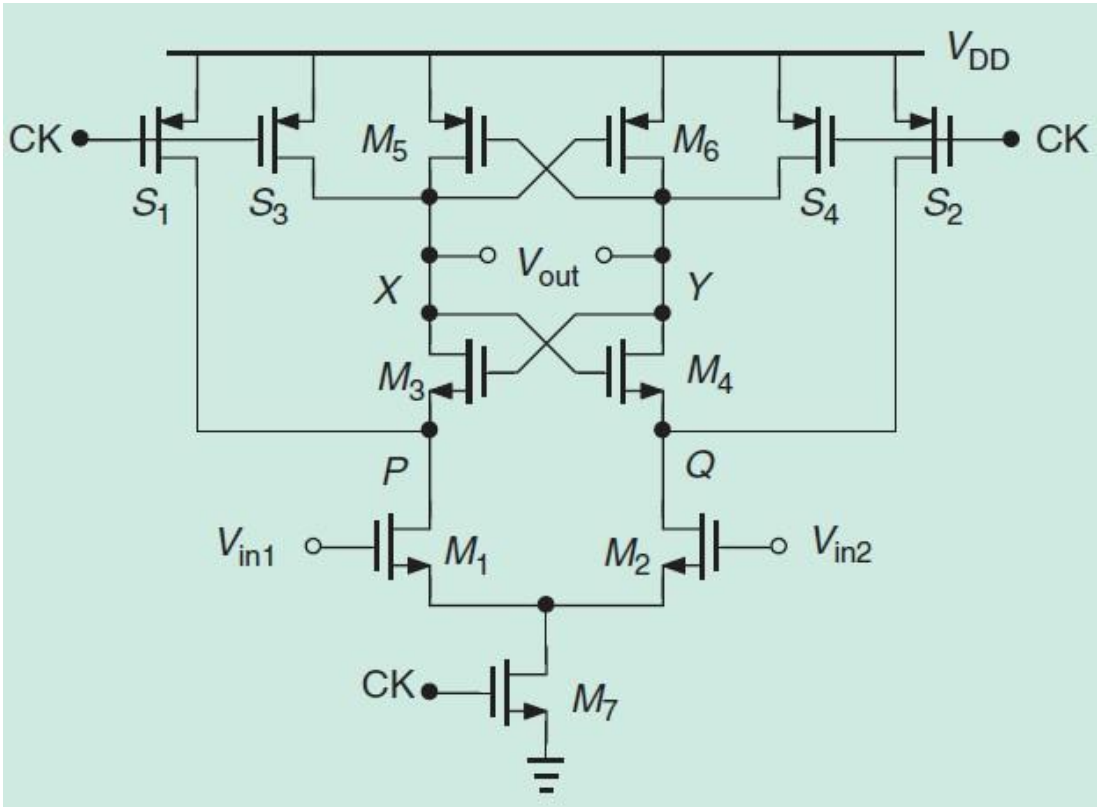
With the above, the offsets from 100 MC samples were analyzed and it was found that the offset variation with input common mode is significantly weakened. 25 samples were chosen at random and their offsets are plotted as a function of the ICM as done before:



The offset variation across ICMs is restricted to about 10mv or lesser this time. The previous SA with its leakage issues did see as much as 60mv of offset variations across the 50mv to 550mv range of ICM. The sizing of all the common devices is kept the same as the previous SA, thus verifying that even a few millivolts difference between Q and Q_bar while in reset can drastically alter SA performance during the amplification stage. The Q and Q_bar difference with the PMOS header in place is restricted to just a few microvolts this time.

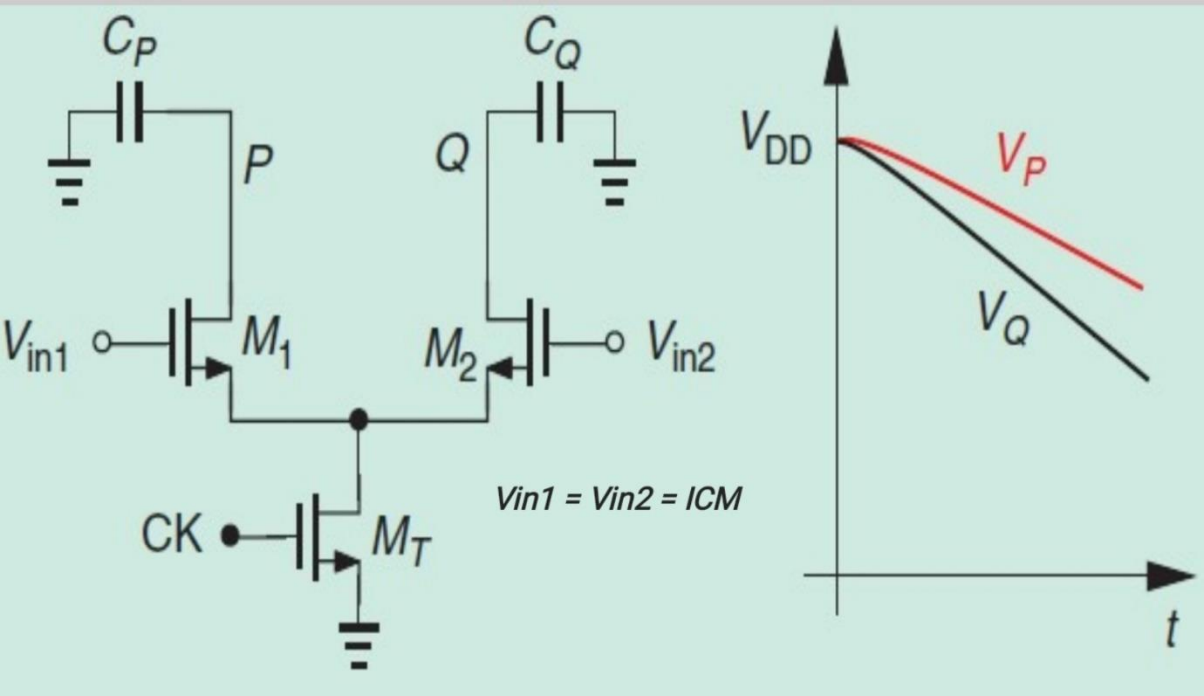
2.3 SIZING THE STRONGARM LATCH

We work with the NMOS differential pair-based Strongarm latch. It doesn't have static leakage issues during the reset phase, which is essential to not only energy efficiency but also resolution performance as seen in section 2.2. We go about sizing the latch to reduce offset by analyzing the MC waveforms and tweaking device sizes. The strongarm latch has the following circuit:



A big advantage of using this structure over the cross-coupled latch is that the strongarm requires just a single clock phase for its operation and this makes the control along with timing much simpler at the chip level. First, we size the devices in this latch assuming a single finger for every device to get good offset behavior. Then, we adjust the number of fingers with each finger at this newly determined size to get the overall offset within desired specifications.

To reduce offset, it is imperative to understand what creates offset in the Strongarm latch. Offset is composed of two components- static and dynamic offsets. The static offset is due to device level mismatches between pairs of transistors that are supposed to be perfectly matched in principle (like M1, M2, or M3, M4, etc. in the figure above). This can be modeled as a fixed input-referred DC voltage source alongside the ideal latch. The second component is the dynamic offset which is primarily due to capacitive mismatches between nodes that are supposed to see identical capacitive loads (like P, Q or X, Y, etc.).



Consider the circuit above where the differential pair begins to discharge P and Q. The cross-connected inverters have not been triggered here yet. The inputs are shorted as above and an appropriate high enough input is applied. The differential rate of discharge between nodes P and Q is a function of not only the capacitive mismatch between P and Q but also the average discharge current. This discharge current will be a strong function of the ICM level as it is applied directly to the gates of M1 and M2. Thus, the offset due to capacitive mismatch is also partly dependent on the input common mode quantity for the Strongarm latch. And the true extent of dynamic offset degradation can be gauged from the layout because improper layout can worsen capacitive matching between nodes and affect overall offset. We attempt to reduce these circuit-level mismatches by increasing the sizes of the devices. The reasoning here is $\sigma_{vt} \propto 1/\sqrt{(WL)}$. Making the critical devices in the Strongarm larger would mean lesser mismatch

between the supposedly matched pairs and hence help reduce the offset variance or spread. Note that W and L refer to the width and length of the transistor respectively.

2.3.1 SIZING METHODOLOGY

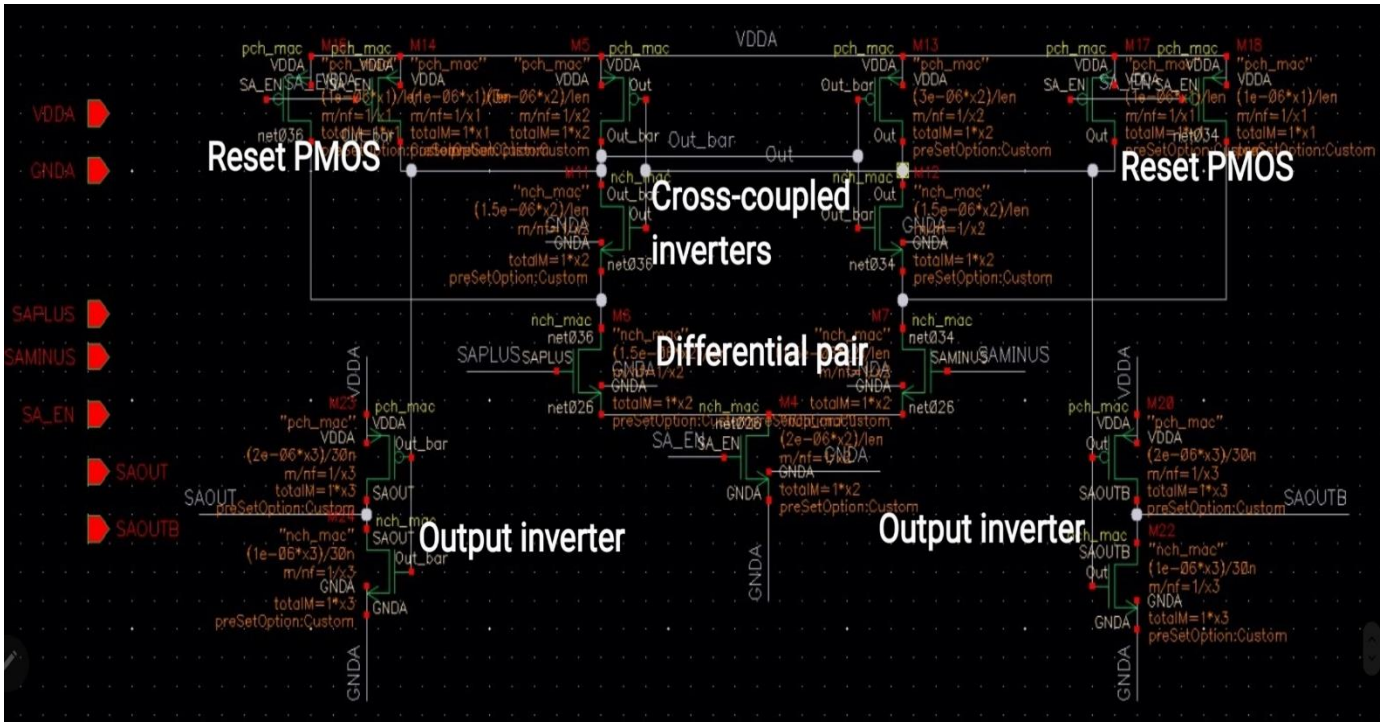
To size the latch, we will repeat the MC offset tests where we step the input differentials and ICM to observe SA transitions for fetching offset values. The devices' sizes are made design variables and iterated upon to optimize offset by inspection of the MC waveform. Once the device sizes are ascertained, we set the number of fingers as design variables and iterate over them to meet specifications. This is verified by automating the offset extraction with a simple Python script that outputs the offset histogram for desired sizes. Following were the observations made while sizing a single finger (i.e., all devices kept at a single finger) for the Strongarm latch-

- Best offset is when the PMOS width for the cross-coupled inverters is kept at maximum i.e., 3um. And the respective NMOS' are roughly half that width and kept here at 1.5um. The NMOS inside the differential pair should have widths that are close to that of the cross-coupled inverters' NMOS and also kept at 1.5um. The tail current NMOS should be able to deliver the desired current for quick resolution and it helps to have a width greater than that of the NMOS' in the differential pair. It is kept at 3um here.

- All devices are kept at minimum length of 30nm. This is done assuming an area-constrained case. We know that the area (W*L) must be increased to reduce mismatch but there exists a second-order effect when we compare Strongarm latches with fixed area. When W*L is fixed and L is increased, we reduce W, and subsequently, the drive strength of all the devices in the latch takes a hit as W/L has reduced significantly. The weaker currents not only take longer to discharge the nodes and produce the outputs but are also more vulnerable to noise and mismatch inaccuracies. The MC output conforms to this prediction as offsets worsen at constant area with increasing lengths. The short-channel effects in reduced length devices didn't seem to compromise offset performance appreciably as long as the short-channel effects are closely matched on either device owing to the large area. Hence for the given area, 30nm length works best.

- The reset PMOS' S1 to S4 are kept at 1um width and minimum length. These shouldn't be made too small as they might then fail to pre-charge the internal nodes to VDD and

compromise the reset phase which affects offsets as seen. There are two output inverters with PMOS and NMOS widths kept at 2um, 1um respectively with minimum lengths:



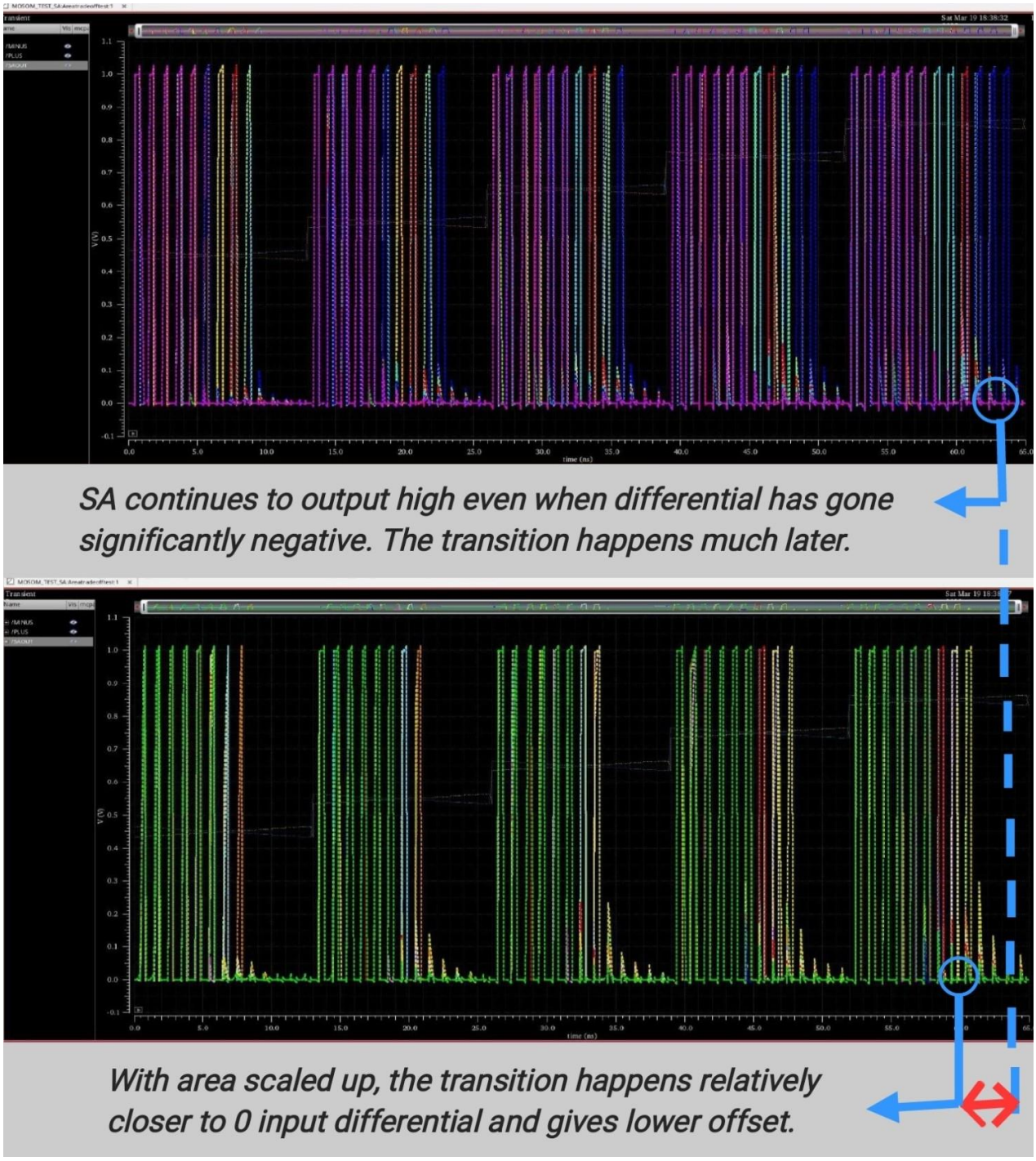
Above is the Strongarm comparator schematic where we have kept the individual sizes and later, the number of fingers as design variables and iterated over them to attain well-behaved offset characteristics.

With these determined sizes, we now adjust the number of fingers of every device to get good offset distribution across corners. We are currently targeting a 4-bit flash ADC with a 0-320mv input range. So, the V_{LSB} = 320mv/2^4 or 20mv. The offset for the comparator should preferably be 0.5V_{LSB}, 10mv at worst. We try to get our SA offsets within ±10mv by adjusting the fingers. With these sizes, the Strongarm latch has offset between -25mv to +25mv for the single finger.

Before we start with the multi-finger test, a few other tests were conducted to verify the area claim for reducing mismatches and thereby offsets. Since the widths are fixed as determined above, in the below tests, the W being scaled n (>1) times is essentially increasing the number of fingers by a factor of n with the L (common to all devices) being kept programmable as well:

1. Case 1 of W, L being scaled n (2 here) times vs. case 2 of W scaled n^2 times and L unchanged. Case 2 performed better in terms of offset even though both had similar areas but case 2 had its devices at n^2 times the strength (W/L) of case 1.

2. W/L is fixed but L is increased n (3 here) times. Turns out that with smaller L, the offset performance is far worse compared to one with a larger L. This is to be expected as the area is scaled n^2 (>>1) times at constant W/L. The two waveforms from the 100-sample MC offset simulation below attest to the same-



3. A test of which property is more beneficial for better offsets- area or W/L strength. Case 1 of 50nm length with 10 fingers in the latch (i.e., differential pair and cross-coupled inverters) vs. case 2 of 60nm length with 10 fingers in the latch. Case 1 has 1.2

times larger W/L but case 2 has 1.2 times larger area. After MC testing, it was found that case 2 with 1.2 times larger area had slightly better offset distribution than case 1, highlighting that area dominates W/L for good offset given sizing is kept reasonable.

A few things to note at this point:

1. The number of fingers in the reset PMOS also needs to scale up in proportion to the number of fingers in the cross-coupled inverters or differential pair (collectively termed as the 'latch'). This is necessary to allow the upsized internal nodes to pre-charge fully to VDD, which is dependent on the reset PMOS strength. A good ratio found from simulations is that number of fingers in every reset PMOS is half the number of fingers in the latch.
2. The number of fingers in the output inverters is also scaled up in proportion to the number of fingers in the latch. The number of fingers here is kept the same as that of the reset PMOS.

2.3.2 OFFSET PERFORMANCE WITH MULTIPLE FINGERS

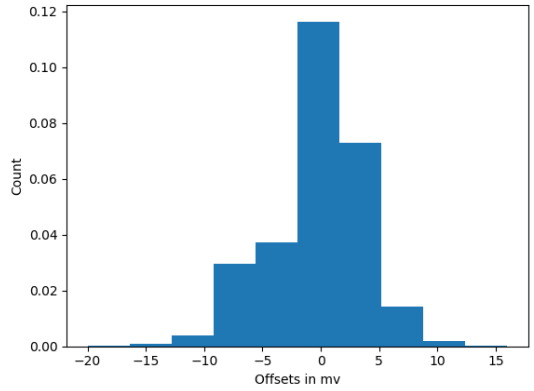
In the present section, the number of fingers signifies the number of fingers in the differential pair as well as the cross-coupled inverters. The number of fingers in the reset PMOS' or output inverters is half the number of fingers in the cross-coupled inverters or the differential pair.

For this test, the input differential was swept from -30mv to +30mv in steps of 2mv and the common mode was swept from 500mv to 900mv in steps of 100mv. The MC output waveform was used as input to a Python script that generated the offset histograms by running a simple logic to capture the SA's transitions. In every offset simulation so far, the reset phase lasts for 400ps and resolution for 600ps. The SA runs at 1 GHz and the timing window used here was found to be more than sufficient by observing the MC output waveforms of the SA. From the constraints of the chip, we have as much as 5ns for the SA to output a decision i.e., we can operate our SA as slow as 200 MHz while still delivering the desired throughput. Offset histograms have been plotted for four cases with a variable number of fingers in the latch as shown below with 150 samples per corner across all corners. Individual finger sizes have been kept as determined in section 2.3.1. The given histograms are generated from the schematic with a noise

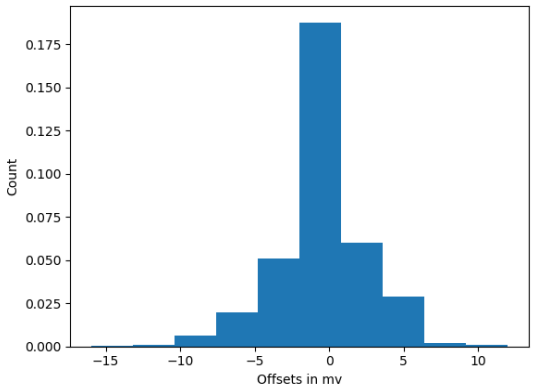
factor of 1 across all the devices, VDD, and ground with a 1 MHz to 10 GHz noise band.

OFFSET HISTOGRAMS

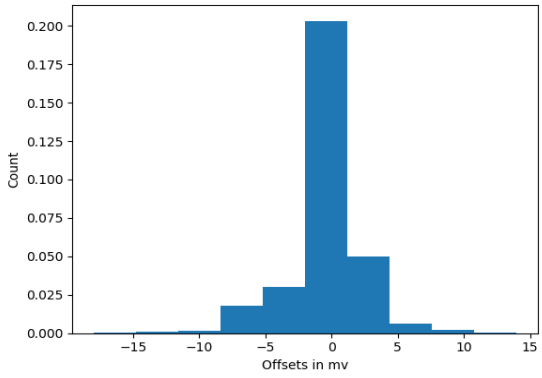
With 15 fingers



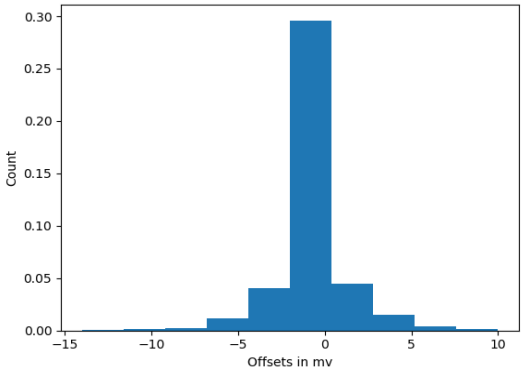
With 20 fingers



With 25 fingers



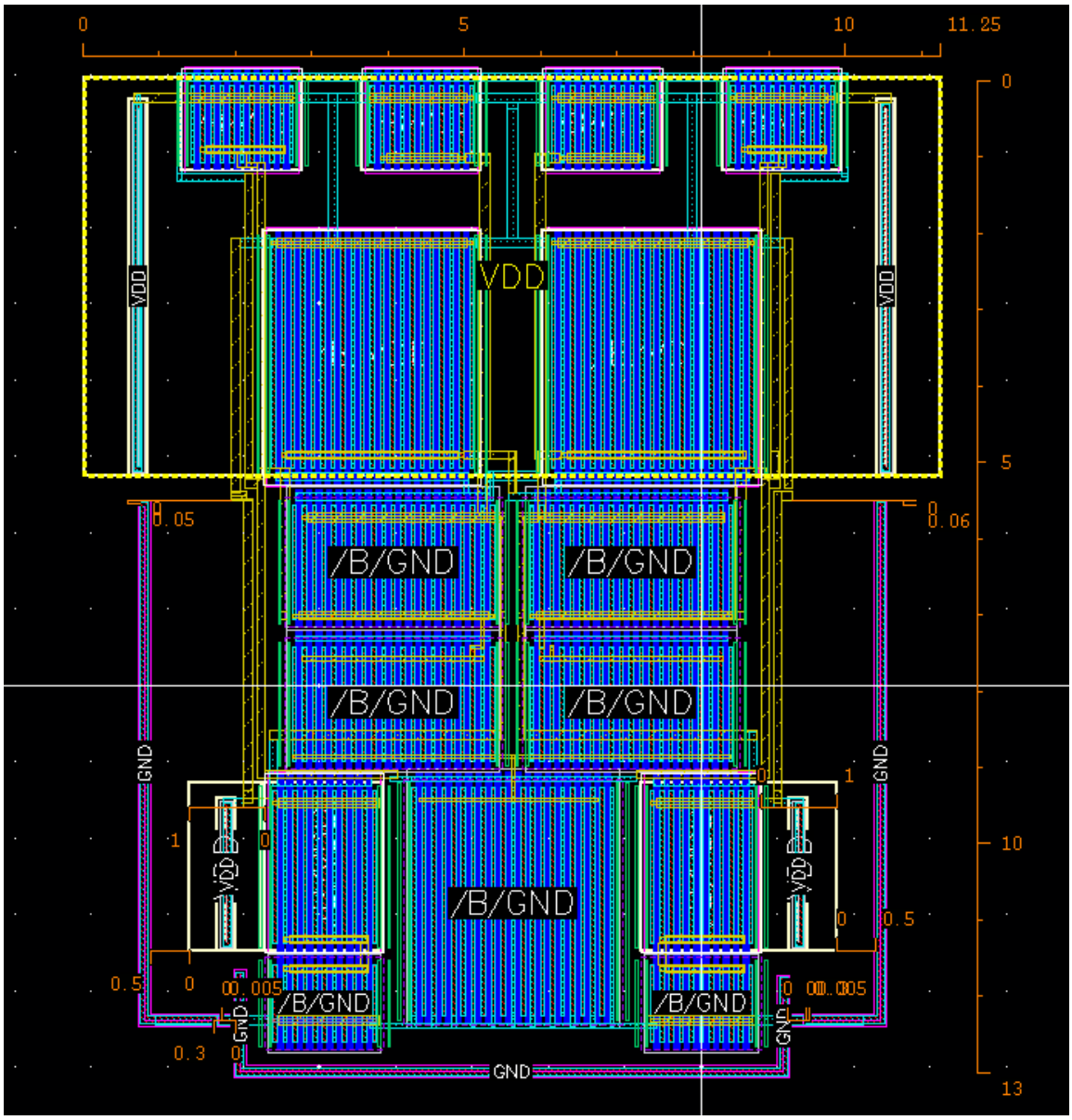
With 30 fingers



As evident, with the increase in the number of fingers, the offset side-bins are increasingly suppressed due to larger areas and lesser mismatches. The SA with 20 fingers in the latch seems to meet our offset specification of 10mv as virtually all samples have offsets within ±10mv.

A DRC-LVS clean layout of the SA with 20 fingers in latch and thereby 10 fingers each in reset PMOS' and output inverters has been designed. Care has been taken to ensure that the

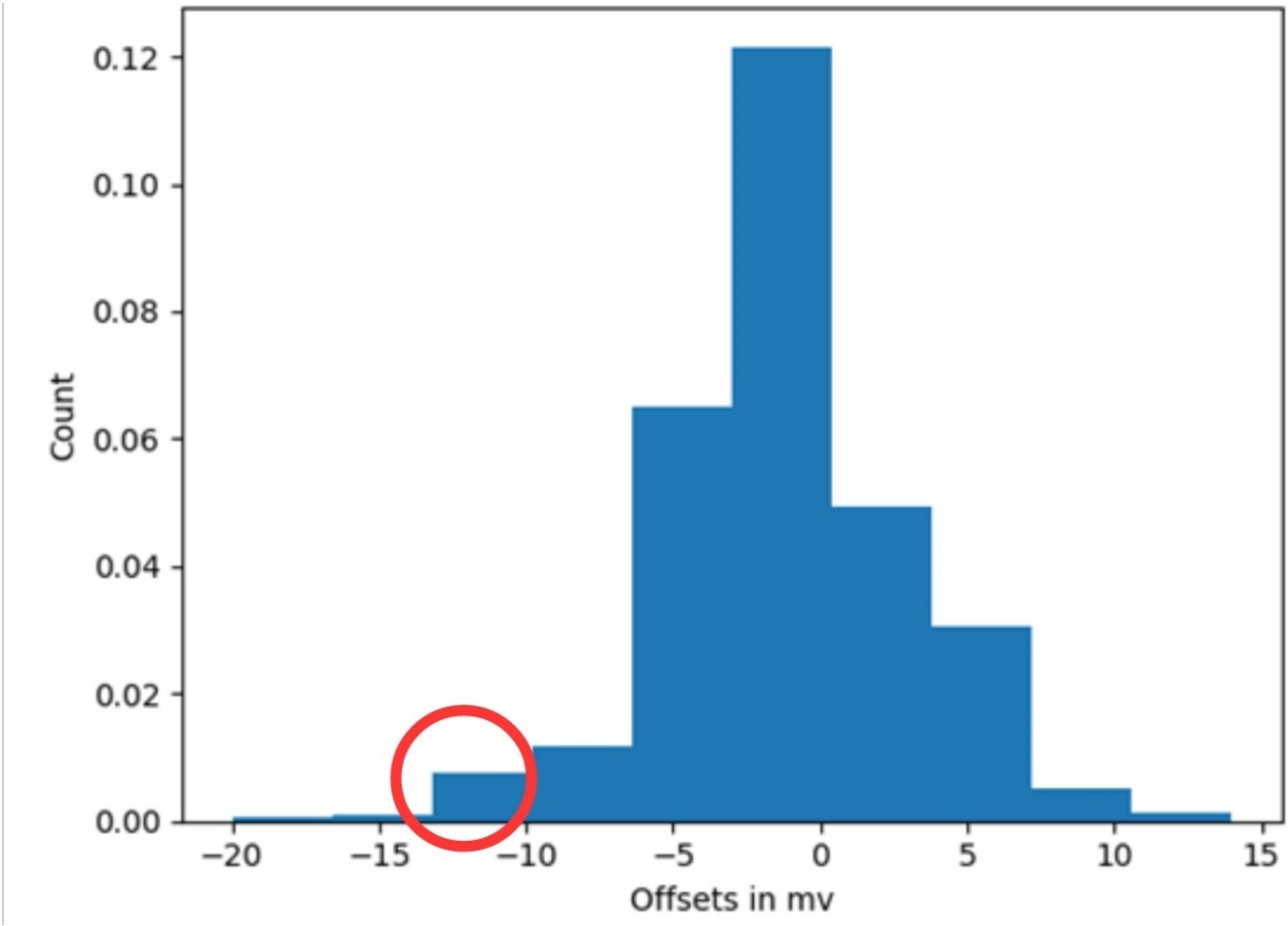
layout is symmetric so as to not induce deliberate dynamic offsets due to capacitive mismatches. The dimensions are 13um * 11.25um.



20 finger SA layout for testing

Parasitic extraction has been done on the above layout (R+C+CC) and MC offset simulations have subsequently been done on the extracted Calibreview to re-check the offset behavior. The ADE L simulations show that the EDP per cycle by integrating power for 1GHz operation of the latch is about 0.22pJ. Peak current consumption of about 3.44mA from the VDD supply is observed during the regeneration phase (very large for a single SA, virtue of enormous W and minimum L). The peak power consumption is, therefore, 3.096mW. An

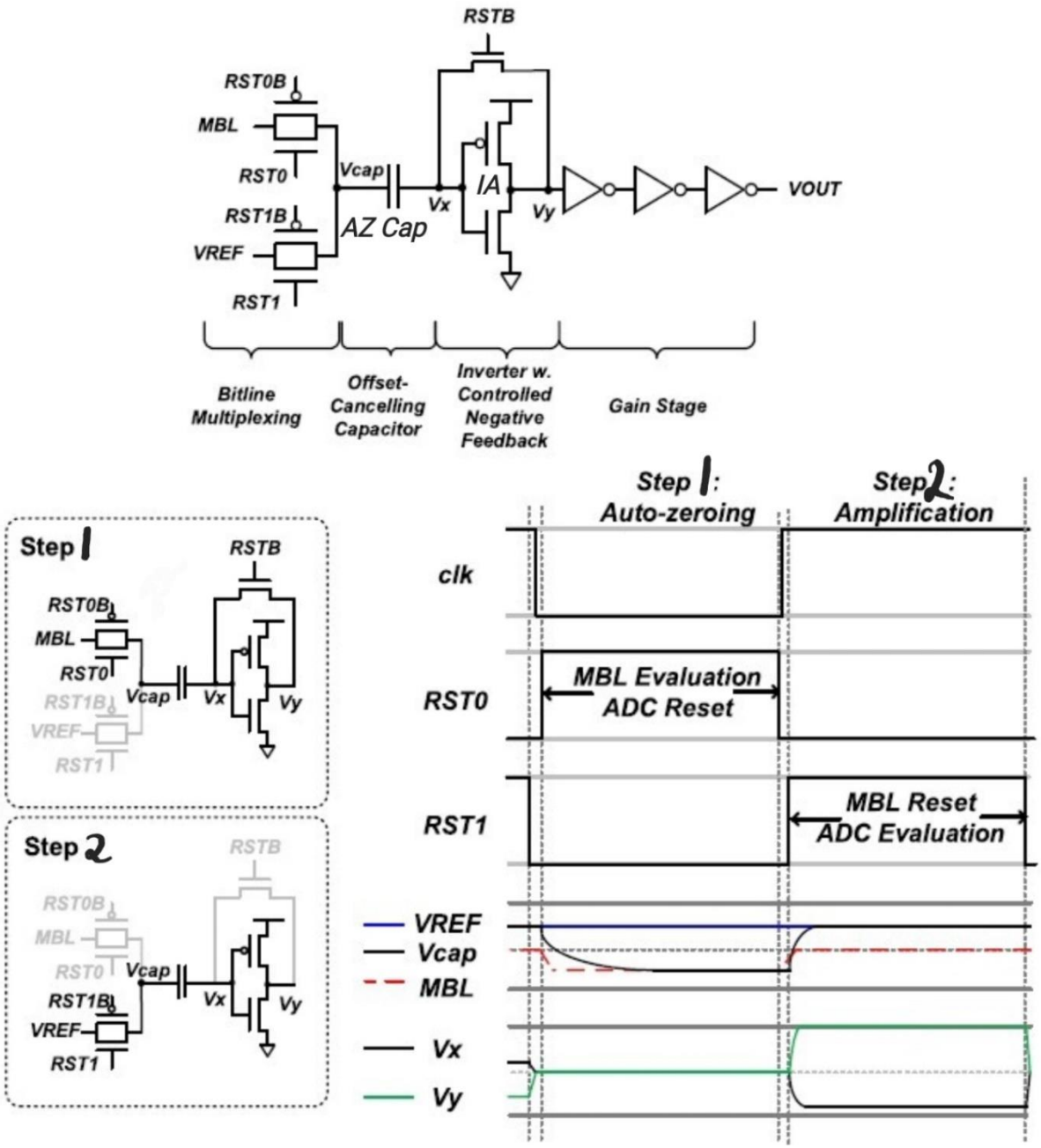
offset histogram has been generated for the 20-finger SA layout after parsing MC output data with the same script. The results are as below.



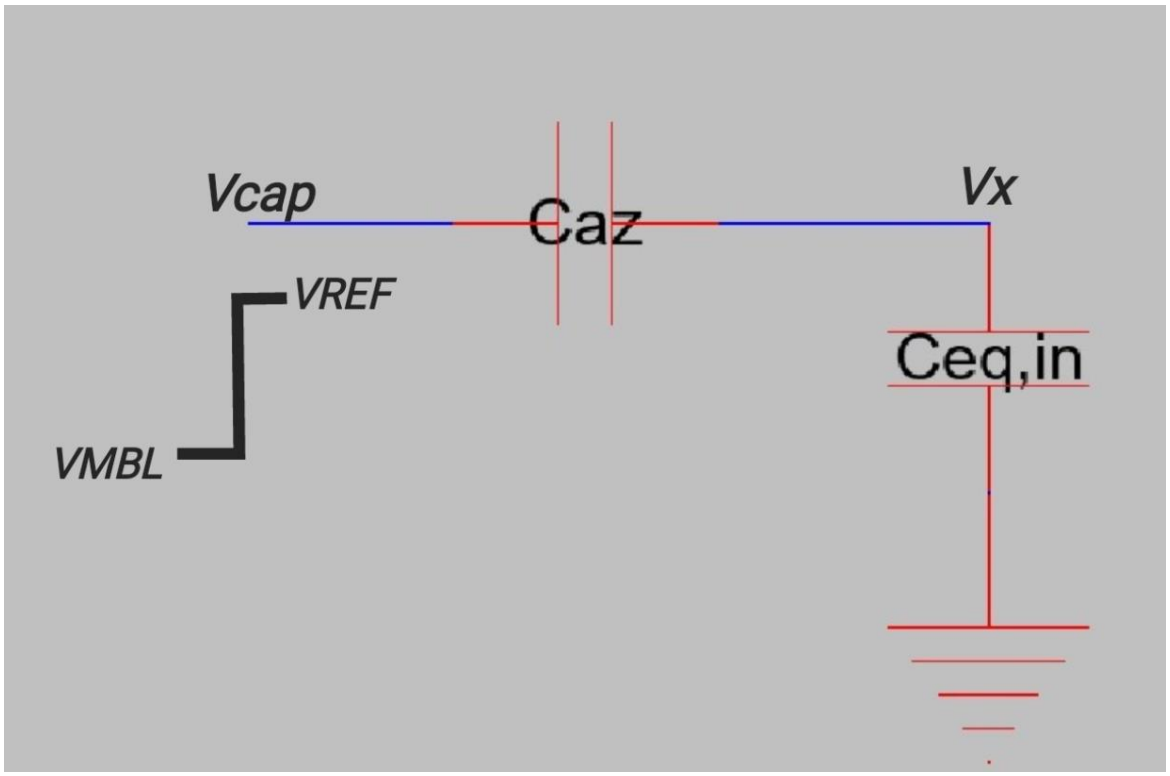
A small fraction of the total samples seem to have their offsets beyond -10mv and this is not strictly conforming with our SA offset specifications seen earlier. So, perhaps the 25 fingers SA might possibly work to produce offsets within desired bounds? This kind of a naïve upsizing strategy may not be the best way of reducing offset as these still rely on matching between devices and nodes which are probabilistic. suffering from area penalty too expensive at chip level. In the next chapter, a new type of SA has been demonstrated that uses a principle of '*Autozeroing*' (AZ) to resolve smaller differentials with enhanced area and energy efficiency. The upcoming design is, as some might argue, simpler than the Strongarm.

Chapter 3 SELF-CALIBRATING INVERTER-BASED SA

3.1 WORKING PRINCIPLE



This SA was proposed in the C3SRAM IMC paper. The operation happens in two stages. First is autozeroing where the trip point information of the inverter IA is stored on the AZ Cap. Here, the RSTB switch turns on and puts the inverter IA in self-bias as its input and output are shorted. Next, we turn on the RST0 switch, and the analog voltage V_{MBL} on the MAC Bitline or MBL is imposed on the node Vcap of the offset canceling capacitor AZ Cap. Then, the MAC evaluation is done on the MBL. Hence, Vx and Vy are driven by IA to its trip point V_M by the end of the AZ stage assuming that the inverter IA can deliver or sink the desired currents. The voltage across the capacitor settles to V_{MBL} - V_M with the Vcap node at positive polarity. The RST1 switch is kept off throughout the Autozero (AZ) duration. The AZ stage ends with RST0 turning off (disconnecting MBL) and then, RSTB turning off (disconnecting self-bias). IA is sensitively balanced and calibrated at its trip point. The second stage is where the SA resolution or amplification takes place. Here, RST1 turns on connecting the reference voltage V_{REF} to the Vcap node. Assuming the AZ Cap to be much larger than the input capacitance of the inverter, the voltage across the AZ Cap can be assumed to stay virtually unchanged as the Vcap node sees a V_{MBL} to V_{REF} transition. Thus, Vx is at roughly $V_{REF} – V_{MBL} + V_{M.}$ If $V_{REF} > V_{MBL}$, Vx goes above trip point and the large gain around the trip point ensures Vy falls significantly below the trip point. This Vy is amplified to digital levels by the inverters that follow it. Similarly, the resolution happens when $V_{REF} < V_{MBL}$. That are a few issues that can hamper the offset performance to be seen.

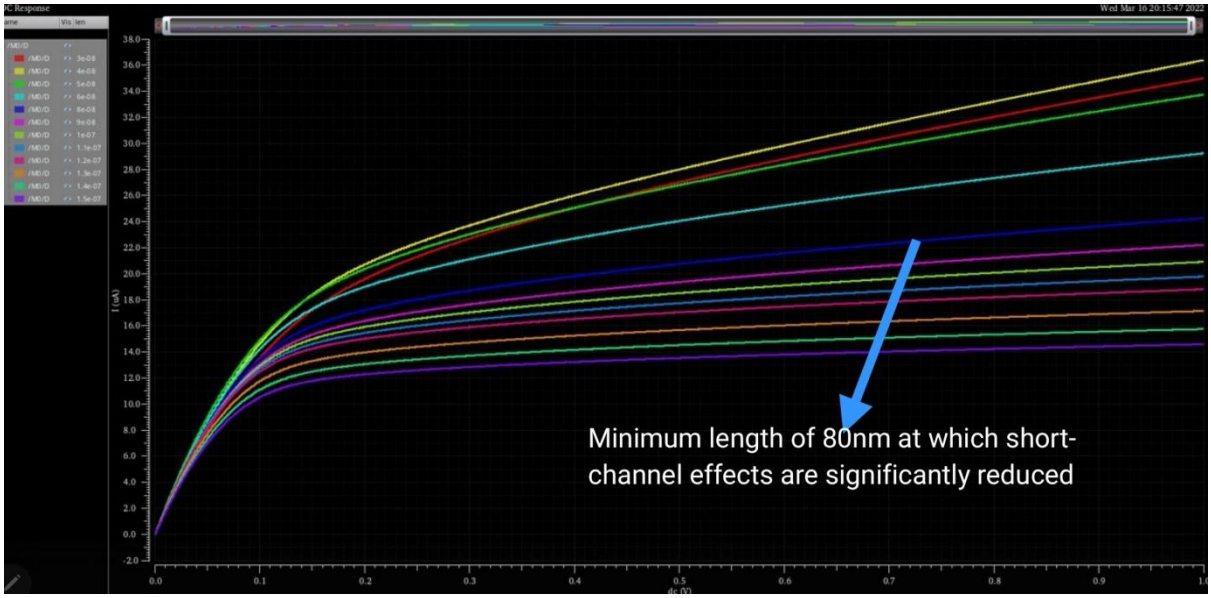


Ceq,in (above) is the equivalent input capacitance of the inverter during the second stage. The entire V_{MBL} to V_{REF} transition can be captured at Vx by making the autozero capacitor Caz >> Ceq,in. This way, less portion of the input differential is lost in capacitive attenuation.

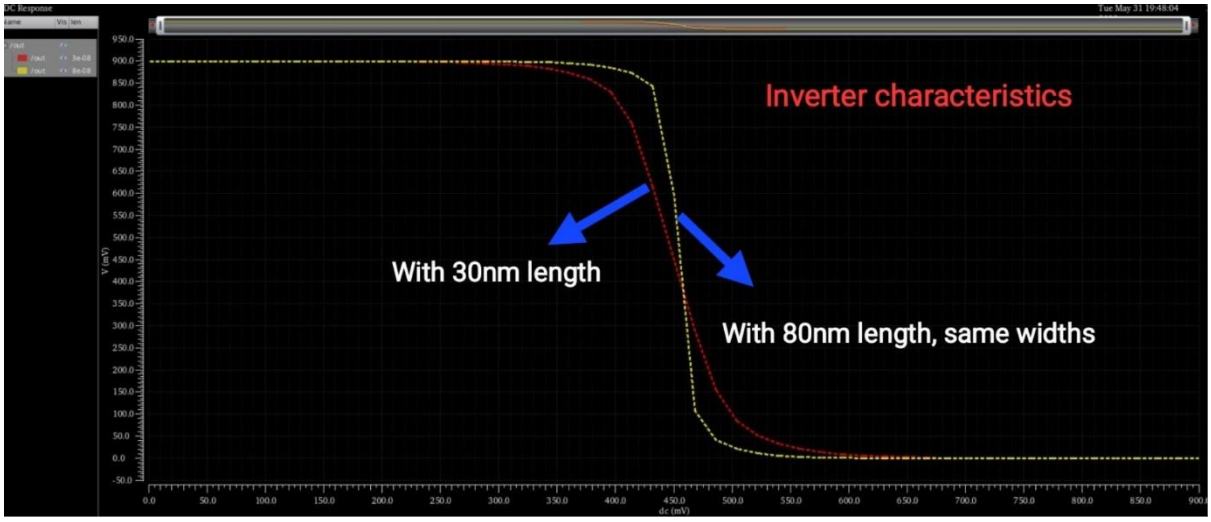
Note that the Autozero capacitor and offset canceling capacitor are the same.

3.2 OBSERVATIONS WITH THE DESIGN

1. The autozeroing inverter IA should have a large gain (Vy/Vx) at its trip point. IA was therefore made with long-channel devices so that short-channel effects do not reduce gain and degrade the SA's amplification. From the Id vs. Vds plots, 80nm was seen to be a reasonable length where the short-channel effects are limited.

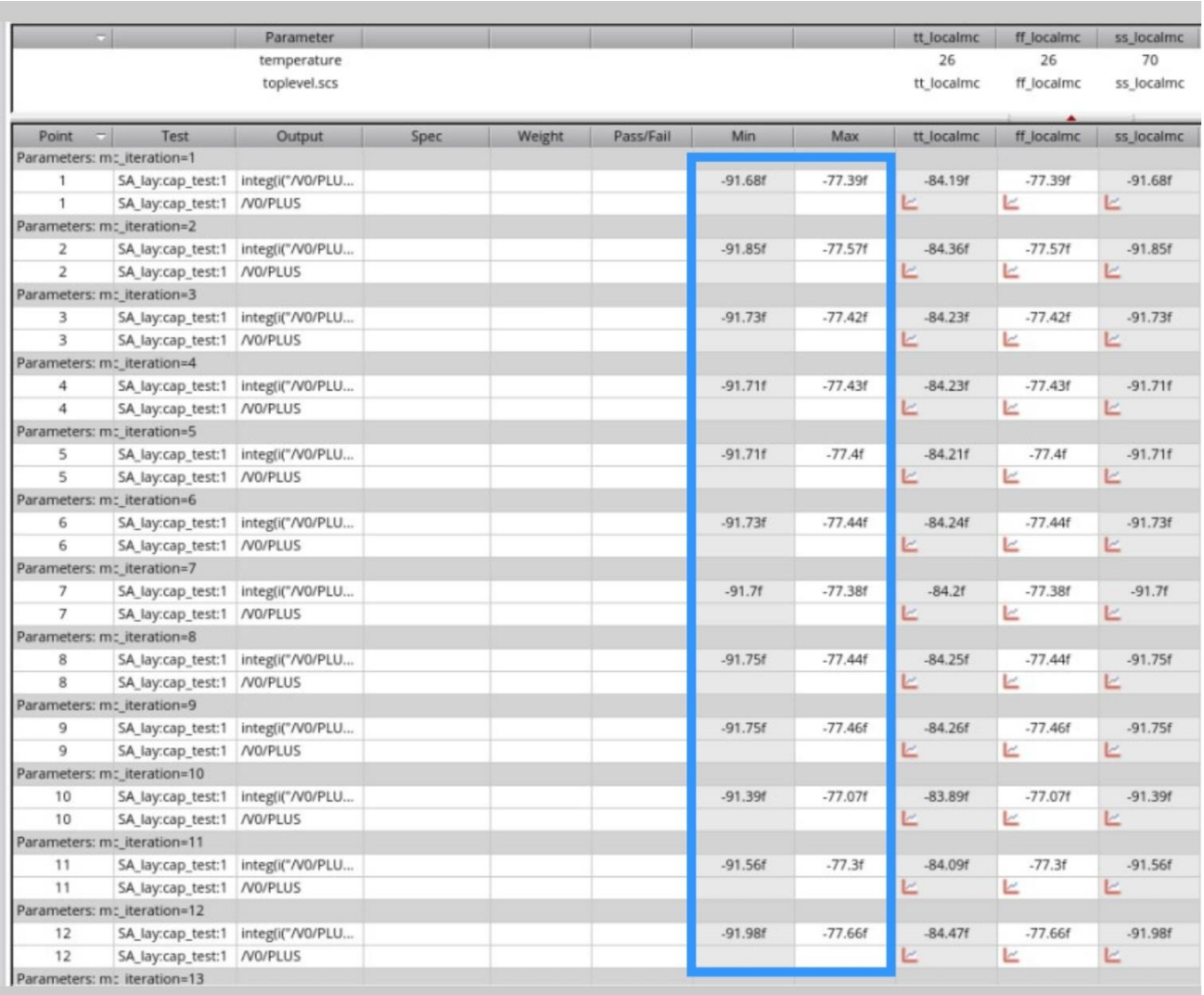


The above exercise was repeated for PMOS and fetches the same result for length. Therefore, the critical sensing inverter IA is made with 80nm length devices with a PMOS width of 3um and NMOS width of 1.6um. The nominal gain at the trip point is around 30, which was found to be high enough to resolve voltages within our desired input differential bounds with sizes of other components fixed at subsequently determined values. The characteristics are compared for steepness when the length is at minimum of 30nm as against the used 80nm. With 80nm, the gain is expectedly higher and more preferable for use in the critical inverter:



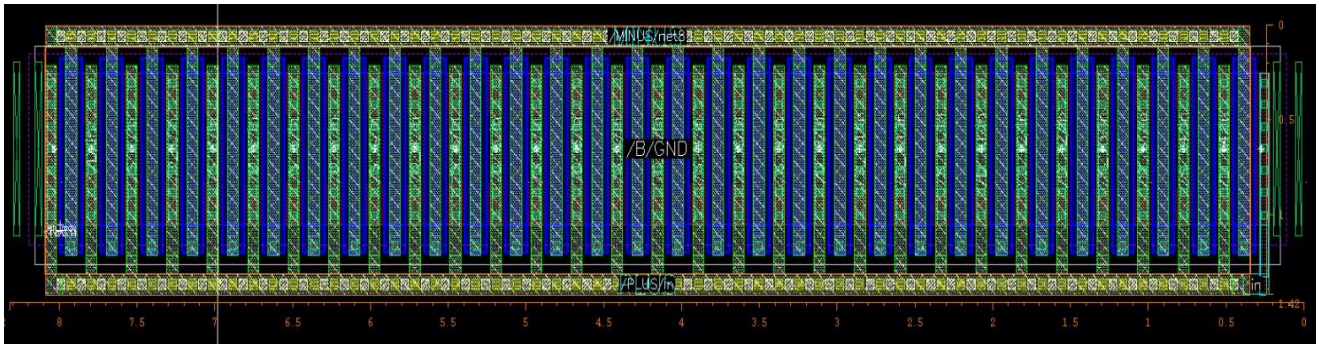
2. The autozero capacitor AZ Cap also called the offset canceling capacitor, is set to 84fF after some considerations. This was fixed from two considerations. From the MC tests of the schematic, it was found that to accurately resolve ±3.5mv input differentials across corners, mismatches, and noise, an AZ Cap of 70fF or above works. The 84fF is realized using 33fF from the MOMCAP cfmom-2t and roughly 51fF from NMOSCAP with both in parallel. This was set from layout considerations as the MOMCAP and MOSCAP areas are roughly equal, and both are overlayed on top of each other for good density. The size of the capacitor was such that the layout was easy to floorplan. The demands from the capacitor are that it should be large enough to combat leakage, input attenuation, and filter away noise on Vx. And it should be small enough for the inverter to autozero correctly sufficiently quicker than 5ns, which is the time available for sensing. The layout of the 84fF capacitor is shown. The MOSCAP contributes around 51fF nominally**, which was ascertained by an independent test to estimate the capacitance. This test involved giving a step transition at the NMOS' gate with its drain and source shorted. The gate current was integrated throughout the step and divided by the magnitude of the step to give the equivalent input capacitance as per the simple equation $C_{eq} = \frac{1}{\Delta V}\int_0^t I_g dt$

**The capacitor was subjected to an MC simulation across temperatures (26˚C to 70˚C) and corners, where the capacitance seemed to vary from 77fF to 92fF, which is beyond the determined 70fF target for 3.5mv offset bound:



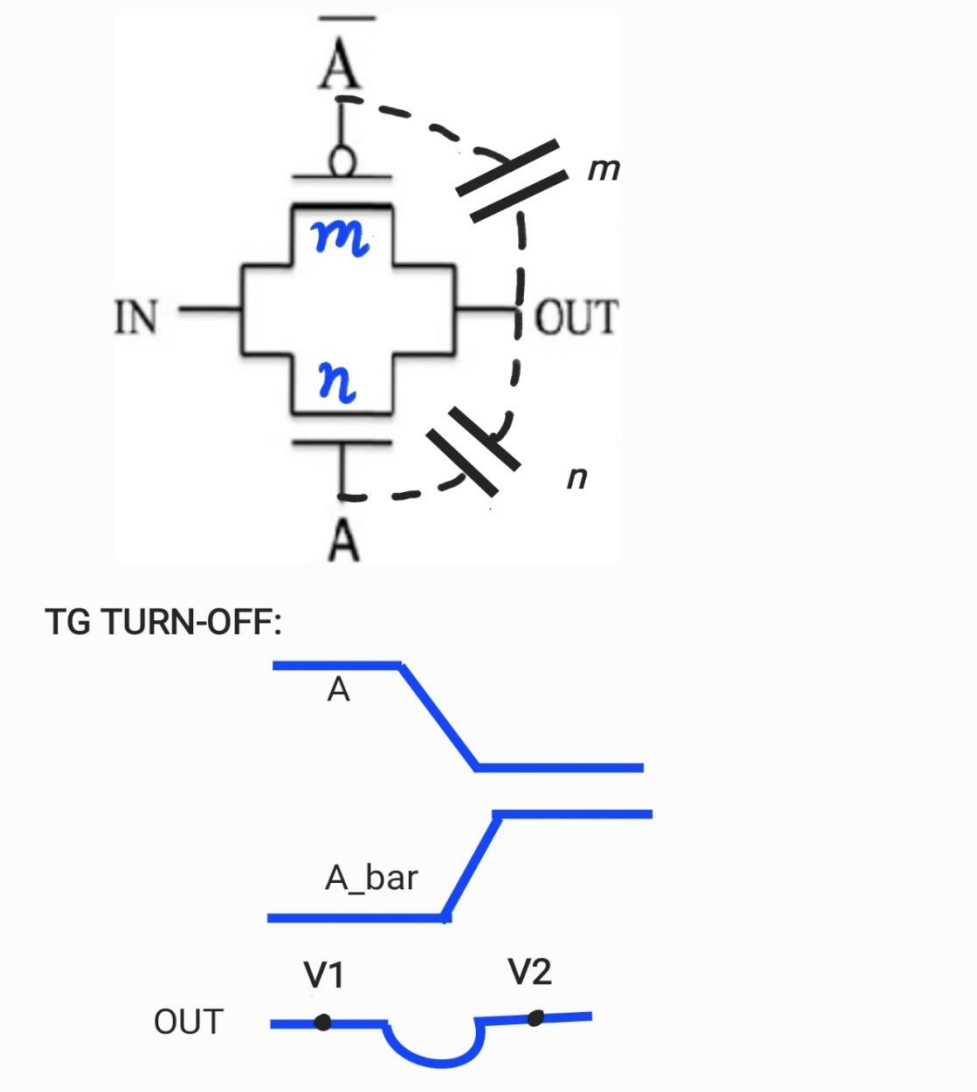
		Parameter						tt_localmc	ff_localmc	ss_localmc
		temperature						26	26	70
		toplevel.scs						tt_localmc	ff_localmc	ss_localmc
Point	Test	Output	Spec	Weight	Pass/Fail	Min	Max	tt_localmc	ff_localmc	ss_localmc
Parameters: m:_iteration=1										
1	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.68f	-77.39f	-84.19f	-77.39f	-91.68f
1	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=2										
2	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.85f	-77.57f	-84.36f	-77.57f	-91.85f
2	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=3										
3	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.73f	-77.42f	-84.23f	-77.42f	-91.73f
3	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=4										
4	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.71f	-77.43f	-84.23f	-77.43f	-91.71f
4	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=5										
5	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.71f	-77.4f	-84.21f	-77.4f	-91.71f
5	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=6										
6	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.73f	-77.44f	-84.24f	-77.44f	-91.73f
6	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=7										
7	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.7f	-77.38f	-84.2f	-77.38f	-91.7f
7	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=8										
8	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.75f	-77.44f	-84.25f	-77.44f	-91.75f
8	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=9										
9	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.75f	-77.46f	-84.26f	-77.46f	-91.75f
9	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=10										
10	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.39f	-77.07f	-83.89f	-77.07f	-91.39f
10	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=11										
11	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.56f	-77.3f	-84.09f	-77.3f	-91.56f
11	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=12										
12	SA_lay:cap_test:1	integ(i("/V0/PLU...				-91.98f	-77.66f	-84.47f	-77.66f	-91.98f
12	SA_lay:cap_test:1	/V0/PLUS								
Parameters: m:_iteration=13										

The effective capacitance of the MOMCAP and MOSCAP in parallel is seen to roughly range from 77fF to 92fF, with a mean at 84fF which is also the nominally determined capacitance of the arrangement. The below layout snapshot shows the dense packing of the MOM-&-MOS CAP, with each occupying roughly the same area and are hence overlaid, within DRC considerations.



The offset canceling ~84fF capacitor layout of size 8.28um * 1.4um.

3. Charge injection on the sensitive Vx node from the transmission gates (TGs) should be at a minimum. We follow a simple static sizing rule here to do the same. We size the PMOS and NMOS in the TG in some ratio m:n such that the net effect of CI from the PMOS and NMOS cancel each other at the output node. The CI is roughly in proportion with the size of the transistors and hence we rely on sizing them appropriately as below:



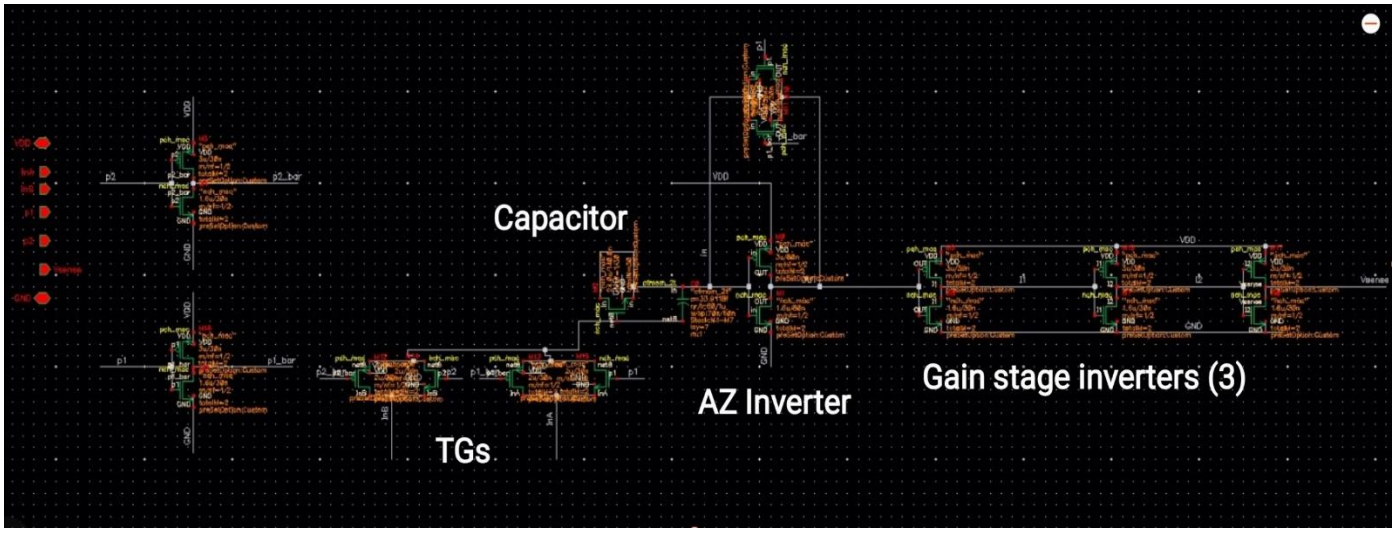
From simulations, we obtain m/n for which V1 ≈ V2 above. It is found that m/n is about 1.05. Therefore, in the TGs, the NMOS and PMOS are kept at 2um widths and minimum length of 30nm for good strength. The inverters other than the autozeroing inverter IA are all kept at a standard size; The PMOS is at 3um width and the NMOS at 1.6um width with lengths at 30nm. Another benefit from this topology of autozeroing structures is that in principle, the offsets of such AZ SAs don't depend on ICM. This is primarily because the sensitive Vx node only sees the input differential applied to it. Vx is at trip point during AZ and is superimposed with the input differential during the amplification stage. The common mode levels do not concern the critical node Vx. Hence, the issue of misbehaving dynamic offset does not arise for this SA.

Now the performance of the SA is also particularly critical to the timing scheme of the control signals. A few things to note in the timing control for the SA, verified by ADE L runs:

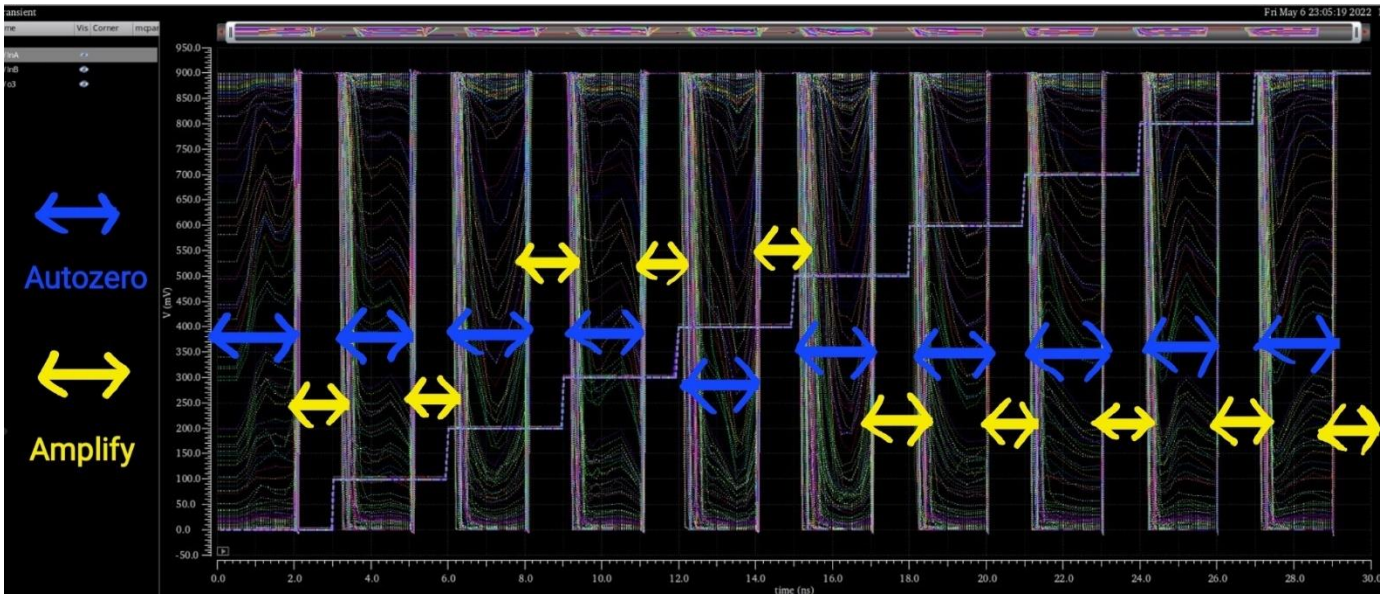
a. Disconnect the first input V_{MBL} using RST0 before disconnecting the self-bias using RSTB. This is important because having it the other way around would mean the sensitive trip point floating voltage on the Vx node at the end of autozero phase could be disturbed and compromise performance. The MAC evaluation must complete before we disconnect the first input from Vx through RST0.
b. Connect the second input V_{REF} using RST1 only after the self-bias has completely shut off using RSTB. This is necessary as otherwise; the AZ capacitor can discharge through the path into the inverter. In other words, the information pertaining to the first input may be partially lost and performance may take a hit, particularly when the input differentials are small.

3.3 SA SCHEMATIC SIMULATION RESULTS

Following is the schematic of the autozeroing SA for which we test the performance:



Note that RSTB used here is a TG and not a single NMOS device as shown in section 3.1. This was done to allow the switch to fully conduct the trip point voltage and also to reduce CI as seen earlier. RST0 and RSTB have been shorted into a single control p1. The layout has been done to ensure that p1 turns off the RST0 TG before propagating to RSTB TG. The justification for this order of turn-off was seen in section 3.2. The p2 signal serves the role of RST1. In the above schematic, the p1_bar and p2_bar signals needed for the TGs are locally generated from the two inverters on left. Performance is tested for this schematic @333 MHz, +3.5mv input differentials at varying ICMs spanning 0 to VDD across corners with variations and noise enabled at all applicable instances:



250 samples are tested with 50 across each corner. The AZ phase lasts for 2ns, which was found to be sufficient across all corners as Vx and Vy nodes saturate to the trip point values. The amplify stage lasts for 1ns as all samples for the schematic can resolve within 500ps. The

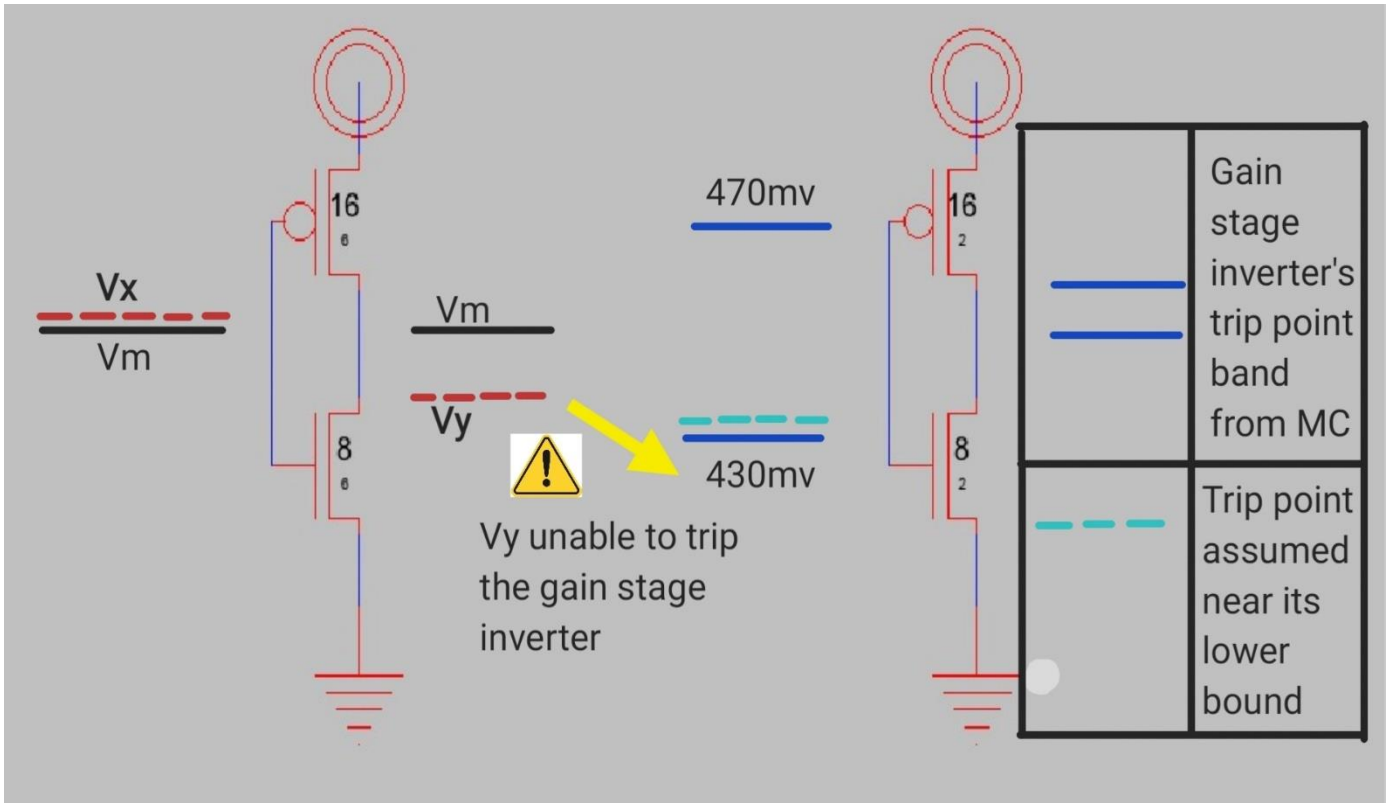
staircase waveform above are the two inputs at 3.5mv differential and the inputs appear to overlap with the ICM being swept from 0 to VDD in 100mv steps. The other waveform is the output of the SA. We see all 250 samples during each amplify phase resolve to high. Observe the low level that remains vacant during amplify verifying that no sample failed to resolve +3.5mv differential at any common mode. The same test when repeated for -3.5mv sees every output sample correctly resolving to a low within the same time. Hence, the offset from the schematic is ≤ 3.5mv. So, the offsets are well within our requirement of ±10mv range.

3.4 SOURCES OF ERROR

Few issues can make the SA output unreliable, particularly at small input differentials and thereby limiting the performance of the SA. These are listed down qualitatively with a schematic simulation result to elucidate this performance limitation during the amplify phase.

1. Capacitive attenuation of the input differential, seen in section 3.1. The equivalent input capacitance of the critical inverter if comparable to the offset canceling capacitor can significantly reduce the magnitude of the transition (or change) seen at the Vx node.
2. Leakage throughout the duration of the amplify phase as the charge stored on the offset canceling capacitor can drop, although slightly owing to junction leakage over time. This and the previously indicated issue are not problems that alone can upset the SA output. This is because these artifacts only attenuate the input differential but do not reverse the polarity of the transition. What creates errors are the following issues that superimpose on the attenuated transition on Vx and swing output the wrong way.
3. The TG's charge injection on the Vx node can restrict performance as well. The sizing of the PMOS and NMOS in the TG as done before helps reduce this effect. But due to the discrete sizing of the devices, its effect on Vx cannot exactly be made zero. However, a large offset canceling capacitor is less likely to lose its stored information from the CI of its neighboring TGs.
4. The large-sized critical inverter along with the TGs can inject noise on the sensitive Vx node and upset its already reduced transition magnitude transiently. The offset canceling capacitor's role is also to serve as an AC short circuit at the noise frequency and short the Vx node to the fixed V_{REF} which serves as the small-signal ground.
5. Mismatch of trip points between the critical inverter and gain stage inverters. This is an issue where even though the critical inverter's output has swung in the right direction

from its trip point, it is not sufficient to trip the gain stage inverters towards the correct rail. This may be solved by upsizing the gain stage inverters to reduce trip point variance or we may use a cross-coupled latch to sample outputs from the critical inverter at the end of each phase and use the latch to compare these sampled outputs instead of propagating it through the feed-forward amplification chain. This is useful because the *optimally sized* single finger of the simple cross-coupled latch has an absolute offset range (25mv) that is smaller than the trip point spread (40mv) of the designed gain stage inverters. The optimally sized 1-finger latch referred to here has PMOS width of 3um, NMOS width of 1.5um with minimum length of 30nm, following a procedure seen for the Strongarm latch in section 2.3. But with this cross-coupled latch, we will require TGs to sample both its inputs before enabling.



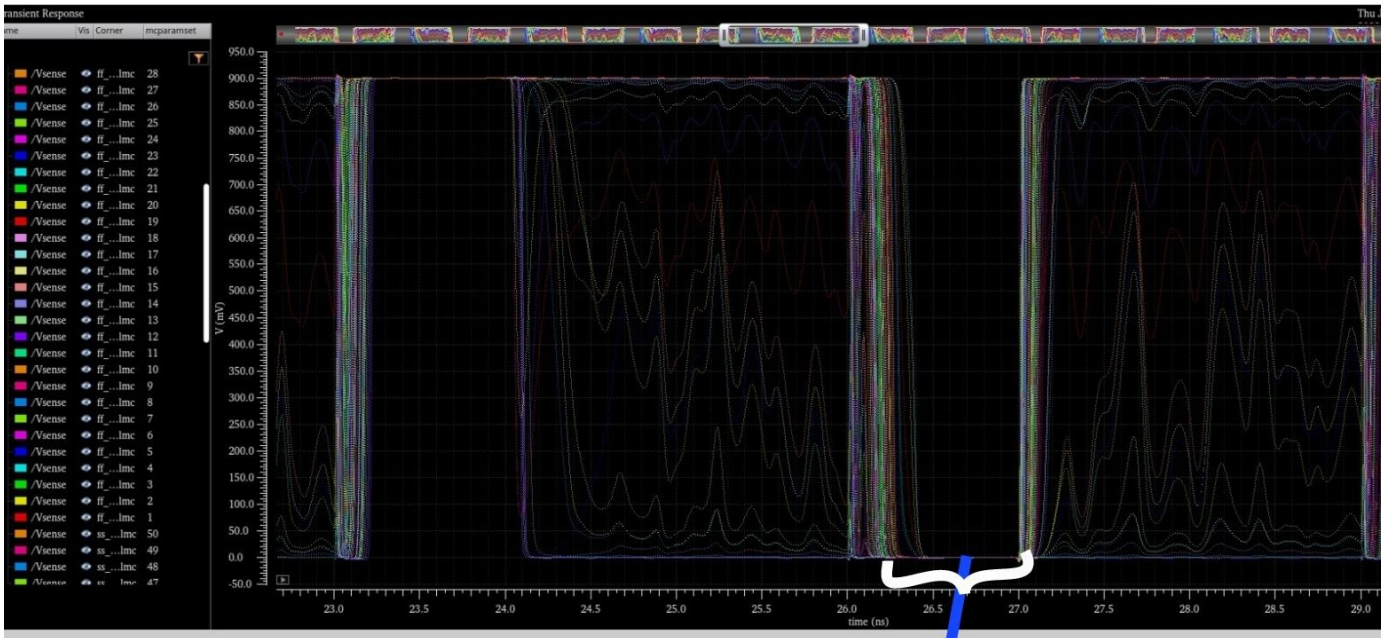
A sufficiently large capacitor solves 1, 2, 3, and 4 seen above. We will subsequently see the capacitance bounds at which the reduction of input transition on the Vx node is made negligible. But for layout, we work with an 84fF capacitor which meets our targets reasonably.

A test was conducted to test the offset limitations of this SA on a nominal basis. It was predicted from hand calculations that the 3mv input differential is likely the offset bound where a very high yield can be obtained. Additionally, input differentials slightly lower than 3mv would see few samples failing to resolve correctly. This 3mv bound was estimated by individually

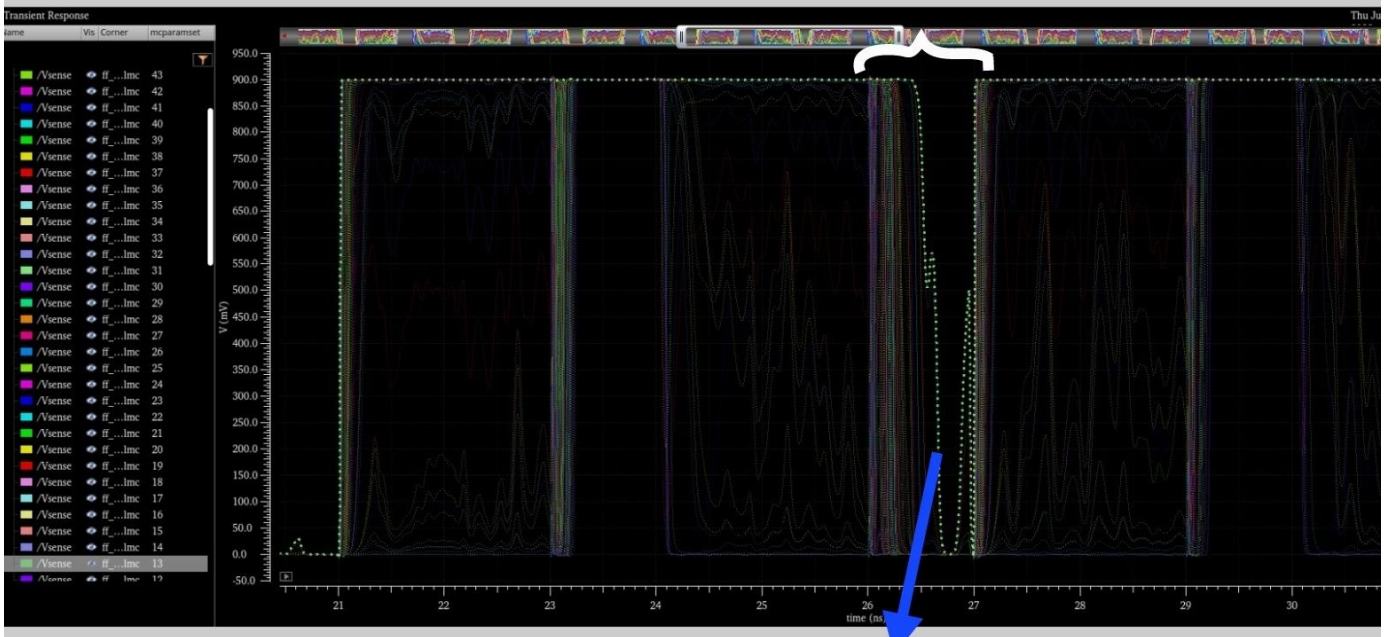
inspecting the sources of error and documenting how much they reduce the input differential that is to be sampled on the node Vx.

Note that the Autozero and Amplify phases have a 40ps non-overlap duration in simulations.

Following the method of section 3.2, the input capacitance was estimated as 5.9fF for the size of the critical inverter used. The test involved giving a steep ramp on the gates as seen earlier for the MOSCAP in section 3.2. The offset canceling capacitor is 84fF. A 3mv transition would therefore be attenuated to 3mv*84/(5.9+84) i.e., 2.8mv. The gate current was integrated due towards the end of AZ phase till the end of the amplify phase and divided by the magnitude of the offset canceling capacitor to fetch the potential that leaked away. It is estimated to be 1.5mv i.e., 0.8mv around when AZ is turned off and an additional 0.7mv during the duration of the amplify phase. The TG's CI diminishes the Vx transition by an additional 0.3mv. The peak-to-peak noise on node Vx is about 0.5mv. Thus, the effective transition seen on Vx in mv is about 2.8 – (1.5 + 0.5 +0.3). This is 0.5mv. But at the end of the AZ phase, the same transition seen on mv is about 2.8 – (0.8 + 0.5 +0.3). This is 1.2mv. The additional 0.7mv dip will occur by the end of amplify phase and was not included here. The Vy node will swing by 30 * 1.2 or 36mv, approximately. The trip point of the gain stage inverters is in a 40mv band, as seen above. This is barely enough to trip the inverter for an extreme corner case seen in the above figure. Thus, 3mv is roughly estimated as the offset bound from the schematic. We might expect some MC samples might misbehave for differentials within the ±3mv range. This is verified through the MC testing conducted below on the schematic with noise enabled. One thing to notice is that we obtain about 3mv offsets from the schematic, although we targeted 3.5mv. This can be credited to overdesigning the capacitor to a nominal value of 84fF, instead of the earlier determined bound of 70fF. Similarly, the capacitor might be increased further for still lower offsets requirements. So, the two ways to increase performance may be to use the cross-coupled latch in place of amplifying buffers and to increase the offset canceling or autozeroing capacitance. But the latency might take a hit as the larger capacitors need more time to charge during the AZ phase. Hence, there is a likely time and area penalty associated with better resolutions. But the designed schematic is sufficient for our purpose and we begin the layout of this SA in the next section. We test if the layout stays true to sensing 3mv input differentials across ICMs, corners, and variations along with noise.



All 100 samples correctly resolve to low at -3mv input difference during marked amplify phase.

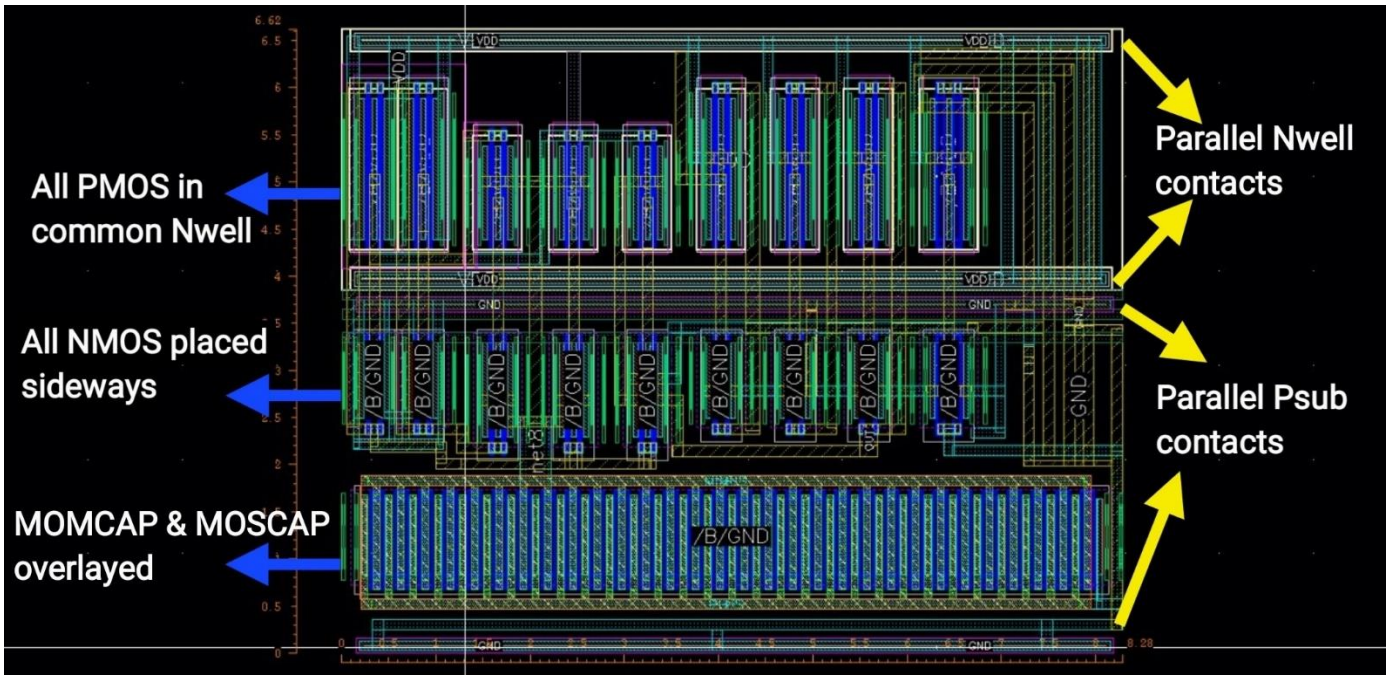


Misbehaving sample at -2.7mv input difference. Doesn't settle to low. Remaining 99 settle to low.

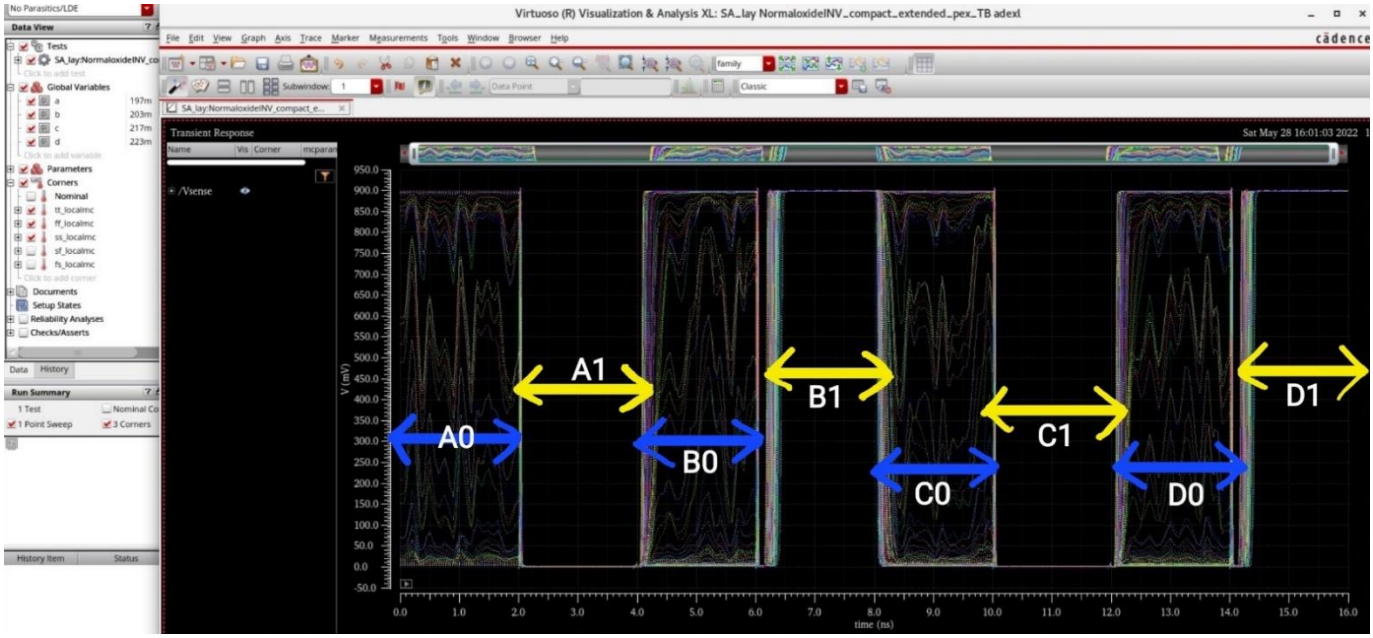
The highlighted failing sample above was attempting to settle correctly to low @2.7mv input difference, but possible leakage during the duration of the amplify phase meant that the diminished transition at Vx is more susceptible to CI and noise. This could result in incorrect sensing, albeit with a low probability as seen here since 1 in 100 samples misbehaved. With this, we can begin arraying the SAs to make the layout for the 15 SAs required to obtain the 15-bit thermometer output code. First, we make the individual SA and test its performance before arraying them for the ADC. We expect similar levels of offsets for this SA array.

3.5 SA LAYOUT AND PERFORMANCE

The layout is height constrained, as we need to pitch match the SA with the DL pitch in the chip. The DL pitch is 6.62um. This will be the height (Y-axis length) of our autozeroing SA. The controls p1 and p2 are propagated from the left and as discussed in 3.2, the RST0 TG is placed more towards the left than RSTB, meaning RSTB's turn off is enabled after p1 has begun turning off RST0. The capacitor is the longest element in our SA at 8.28um. So, the SA layout dimensions are about 6.62um*8.28um. The sensitive Vx node will be shielded by supply rails that run around it in the same metal layer of M3 near the right edge. Strips of N-well and P-sub contacts run parallel to X-axis so avoid latch-up or any form of performance degradation in the transistors. The DRC-LVS clean SA layout is as shown with nearly 73% lesser area than the upsized Strongarm latch of 20 fingers:



The Parasitic Extraction (PEX) for the layout is done and the extracted netlist is simulated to test its performance at ±3mv differentials at various ICMs. As the performance was seen to be independent of ICM, the MC testing of the PEX netlist is done at a roughly constant ICM of 210mv with 50 samples each in the tt, ss, and ff corners. All the 150 samples were seen to correctly resolve the 3mv differentials. Here, the AZ phase lasts 2ns and the amplification phase also lasts 2ns, i.e., a total of 4ns per decision is used. But in the SA array for the Flash ADC, the time taken for a decision would be made 5ns, as dictated from the IMC's MAC duration, with each phase roughly taking up half the decision interval.

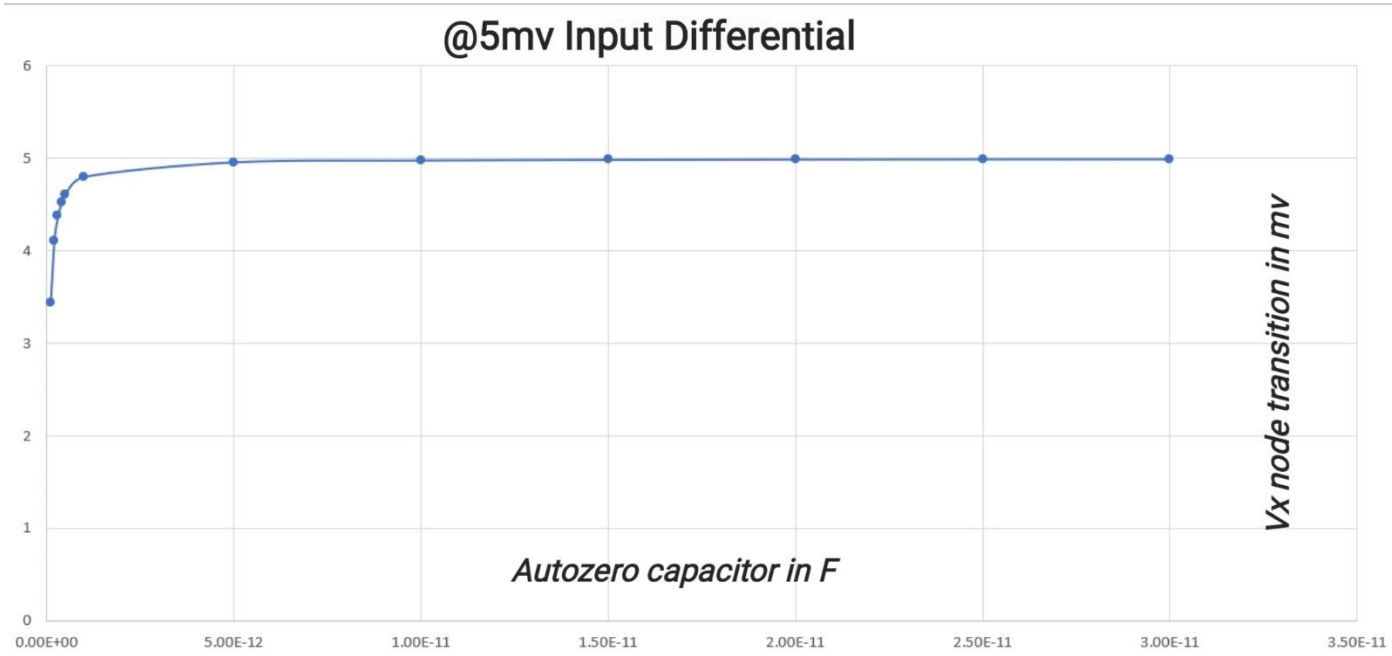


The above A0, B0, C0, and D0 are the autozero intervals that last 2ns each. A1, B1, C1, and D1 are the amplification or evaluation intervals with each lasting 2ns. The plotted waveform is the output of the SA. 150 output samples are evaluated. In A1, 197mv compares against 200mv (X compares against Y means Y connected during AZ phase, X during Amplify) and all samples correctly output low. In B1, 203mv compares against 200mv, and all samples output high. In C1, 217mv compares against 220mv, and all outputs go low. In D1, 223mv is compared against 220mv and all samples go high. This simulation has noise enabled with a factor of one in the 1 MHz to 10 GHz band at all devices, VDD, and ground, as is done in every MC simulation. The experiment may be repeated for arbitrary ICM levels with 3mv differentials and is observed to pass every time. So, the SA layout has offset within 3mv while operating at 250MHz, which comfortably meets our SA specifications.

3.6 OFFSET CAPACITOR BOUND FOR ULTRA-LOW OFFSET

In section 3.4, we had seen ways to improve the offset performance if required later. We saw the problem of input differential being reduced while being sensed at the Vx node. It was qualitatively seen why a large AZ capacitor will limit the reduction of the input differential imposed on Vx. Here, a simple experiment was conducted to figure out what sizes of the capacitor will serve the purpose of better sensing at Vx. A 5mv input differential is applied and the magnitude of the transition is noted at the Vx node as the AZ capacitance is increased. This transition on Vx spans an interval that begins at the end of the AZ phase to the end of the amplify phase (to include effects of leakage even during the amplify phase). Here, AZ and

amplify phases, both last 5ns. This is larger than the timing used for the SA so far since the larger capacitors would require more time for AZ given the same size of the critical inverter and the larger amplify durations assist in leaking more charge from the capacitor, as a worst-case scenario for guiding the design of capacitors. The Vx transition is plotted vs. capacitances:

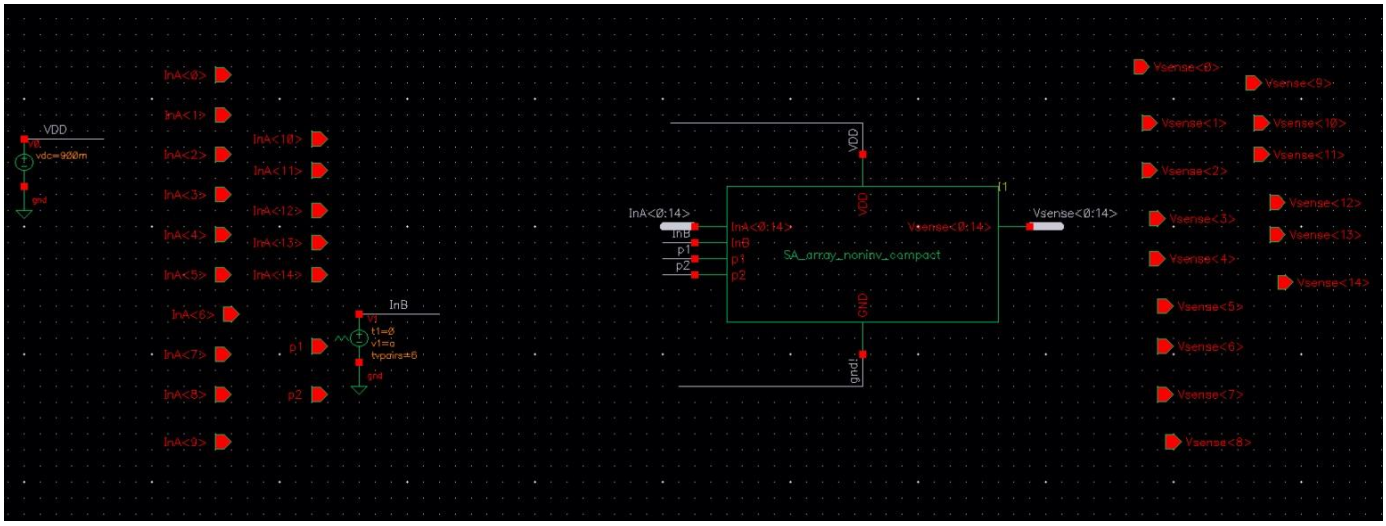


The first point in the above plot is for 100fF (close to the laid out 84fF in our SA layout), where the transition observed at Vx was just 3.44mv, when we ideally expect the complete 5mv. As the capacitor is made 20pF, this Vx transition now becomes 4.99mv, thereby approaching 5mv. The last point on the plot is for 30pF, where the transition is 4.993mv. In theory, to capture the exact 5mv change on Vx, we need an infinitely large AZ capacitor since the entire attenuation phenomenon here can be modeled as a capacitive voltage divider driven by the input differential. As an observation, this voltage divider consists of two capacitors, the first is the AZ capacitor and the second is the capacitances concerning the critical inverter. This second capacitance is roughly the combination of both, the equivalent gate capacitance as well as the 'capacitance' of a model that explains the leakage during the whole sensing process. This second capacitance is roughly estimated as 45fF from the plotted data points, which seems to be considerably higher than the input gate capacitance of 5.9fF mentioned in section 3.4.

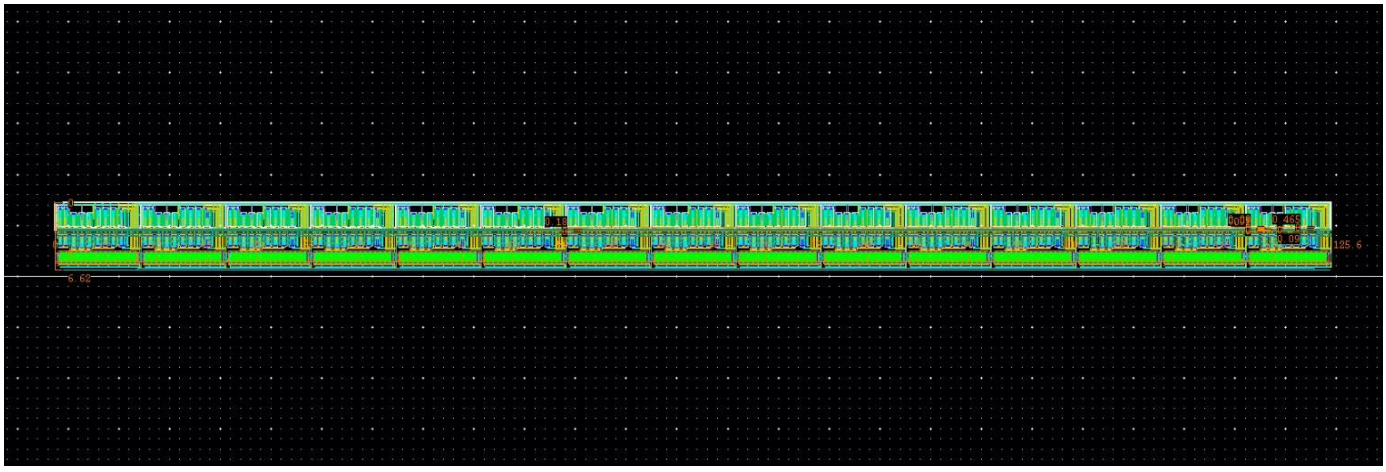
3.7 SA ARRAY LAYOUT AND PERFORMANCE

A 15-SA DRC-LVS clean array was laid out with the dimensions of 6.62um * 125.6um. It is pitch-matched with the DL height. Each of the 15 units is identical to the SA layout in section 3.5. We eventually PEX and perform MC tests to ensure that offsets are within the 3mv bounds.

The following schematic has been realized in layout:

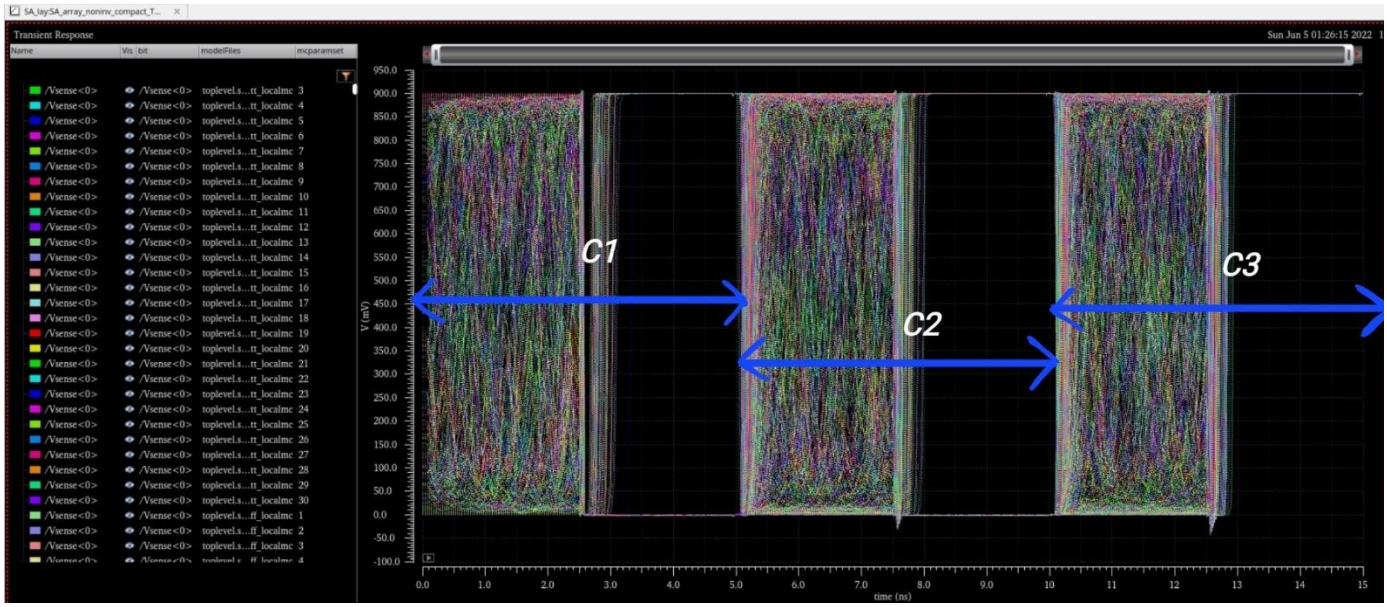


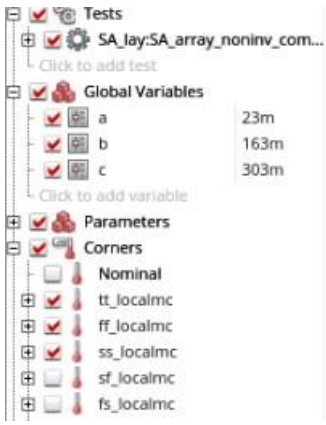
InA<0:14> are the 15-reference voltage (V_{REF}) pins, which are connected to the AZ capacitor during AZ phase. SA_array_noninv_compact is the array of 15 SAs. InB pin is the V_{MBL} common to all SAs, which is the voltage connected to the capacitor during the amplify phase. Vsense<0:14> is the 15-bit thermometer code output to be inspected. The reference voltages to be externally provided are in the 0-320mv range i.e., 20mv, 40mv, all the way to 300mv.



Above are 15 SAs laid horizontally spanning 125.6um. The output is the 15-bit thermometer code with the lesser significant bits being fetched from the SAs towards the left. The InB pin (where V_{MBL} connects) is placed at the center of the layout and not towards any edge to reduce

parasitic attenuation of the InB voltage as traverses the length of the layout. The wire shorting InB across all the SAs is kept 265nm wide which was tested to produce offsets as desired. MC testing is done with the parasitic extracted netlist. The thermometer code is analyzed and verified to be correct for all samples across corners with variations and noise. 30 samples are simulated across tt, ss, and ff corners each, with the ss corner at 70˚C for the slowest performance. Autozero and amplify phases last 2.5ns each and resolution happens well within the given intervals. The obtained waveform is described below.



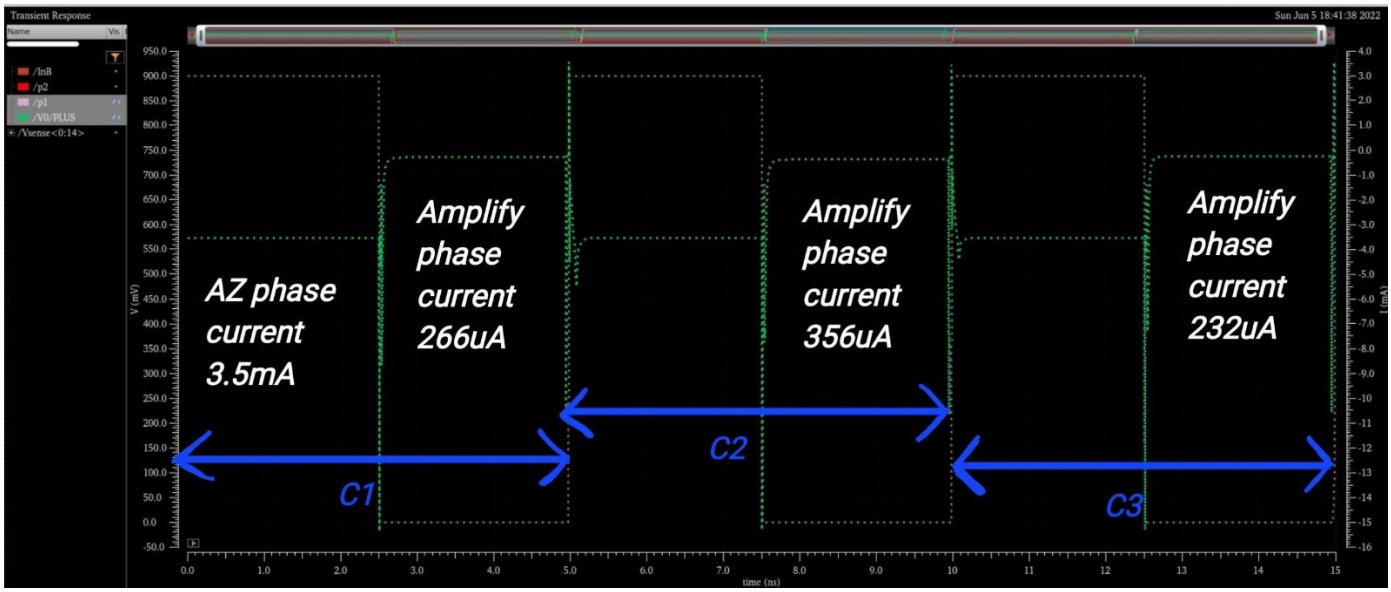


C1, C2, and C3 are conversion cycles where the inputs to the Flash ADC are a, b, and c, which are 23mv, 163m, and 303mv respectively. Each conversion lasts 5ns with AZ and Amplification of 2.5ns each. Plotted waveforms are the outputs Vsense<0:14> of all the 90 samples. It was found that all samples output correctly at every conversion. In C1, where 23mv was the input, the observation was that the Vsense<0> went high, and the rest quickly resolved to low. In C2, where 163mv was applied, Vsense<0:7> resolved to high voltages, and the rest

stayed low. In C3, where 303mv was input, all the Vsense<0:14> signals can be seen to go high. The nominal simulations were also verified to be correct. Similar experiments were repeated with every input to verify the 3mv offset bound expected from the array layout i.e., by giving inputs that are 3mv away from every reference voltage. All the plotted samples were found to output the desired thermometer codes correctly. These experiments were used to conclude that it's best to connect the InB pin at the center of the layout.

Thus, the offsets are well inside the 10mv range required of the ADC. Let us also look at the EDP of the 4-bit Flash ADC. We inspect the VDD current sink and integrate the power over C1, C2, and C3 to obtain the nominal energy consumption over three comparisons. We expect large supply currents during the AZ phase when the AZ inverters have both their PMOS and NMOS in saturation. Due to the AZ inverter's output at the trip point, there may be significant static dissipation in the first few inverters of the gain stage as well during this phase.

Observe the green trace below:



With the same a, b and c inputs seen earlier, we note that the AZ phase current consumption is largely constant at about 3.5mA (Ignore the current spikes as they are likely to be suppressed by On-Chip parasitic inductances). In subsection 2.3.2, it was seen that the 20-finger SA latch alone has a peak current consumption of 3.44mA, almost equivalent to the 3.5mA peak current consumption of the 15-SA array! During the amplify phase, consumption is around the order of a few hundreds of microamps. The energy consumed in 15ns across the three comparisons upon integration is 26.9pJ. The average power consumption is, therefore, 1.8mW. The AZ phase peak power consumption is around 3.15mW. This completes the SA array layout design.

Chapter 4 CONCLUSION

4.1 RECAP

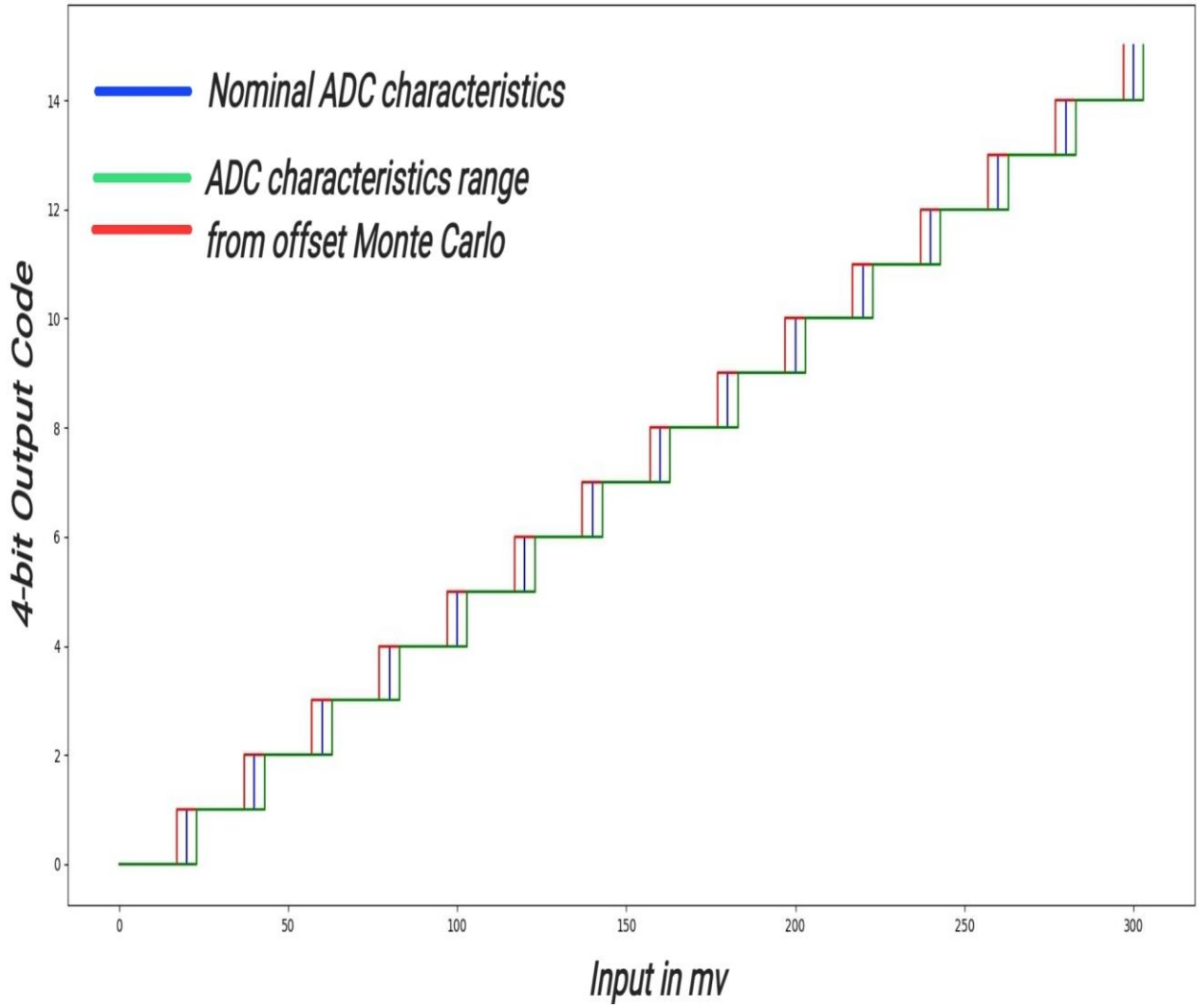
We looked at the Strongarm latch containing the NMOS differential pair and upsized the devices to meet the offset specifications. It seemed an exhaustive way of sizing to meet the offset bounds. The post-layout simulations showed that the offsets had somewhat worsened from the schematic expectations. We needed to overdesign the area significantly to meet the 10mv offset bound. We explored an approach that uses capacitive offset storage as a means of canceling the offset and hence, bypasses the constraints of exact matching between elements of the SA. There are multiple autozeroing structures involving pre-amplifying opamps. The one used here is a simple inverter-based design and was found to be sufficient to meet our demands.

The design of this self-calibrating inverter for autozeroing or offset cancellation forms the heart of the new SA. The timing constraints were briefly outlined and the signal propagation on the layout handles some of those. Simulations were done extensively across corners to verify the output reliability for the same. The use of a cross-coupled latch may be more beneficial for even lower offsets due to the significant mismatches of the AZ inverter with the gain stage inverters. The increased sizing of the offset capacitor to increasingly eliminate input difference attenuation on the gates of the AZ inverter during the amplify phase was also observed through simulations. It was noted that the latency should be increased to allow the larger capacitors to fully charge. The TG's PMOS and NMOS were sized to reduce CI on critical nodes. The layout for the 15-SA array was completed and tested with mismatches, variations, and noise to ensure that the offset specifications are met with good tolerance margins considering a 0-320mv input range for digitization. The capacitor was somewhat overdesigned for this purpose.

4.2 SA ARRAY RESULTS

Offsets were found to be within 3mv for each SA in the array from the layout MC test. The layout is of dimensions 6.62um*125.6um, with each SA occupying 6.62um*8.28um, which is nearly 73% lesser area than the upsized Strongarm latch of size 13um*11.25um that almost had four-fold offset and 15 times the peak current of the inverter-based SA. Highlighted below is the full range of the ADC's input-output characteristics where each SA's offset is at the worst

possible value of ±3mv. Note that the digital back-end has not been explicitly laid out here but its functionality has been assumed to obtain the 4-bit ADC's standalone performance metrics:



The DNL is less than |± (3*2)/20| or, 0.3 @ 3mv peak offset.

The INL is less than |±3/20| or, 0.15 @ 3mv peak offset.

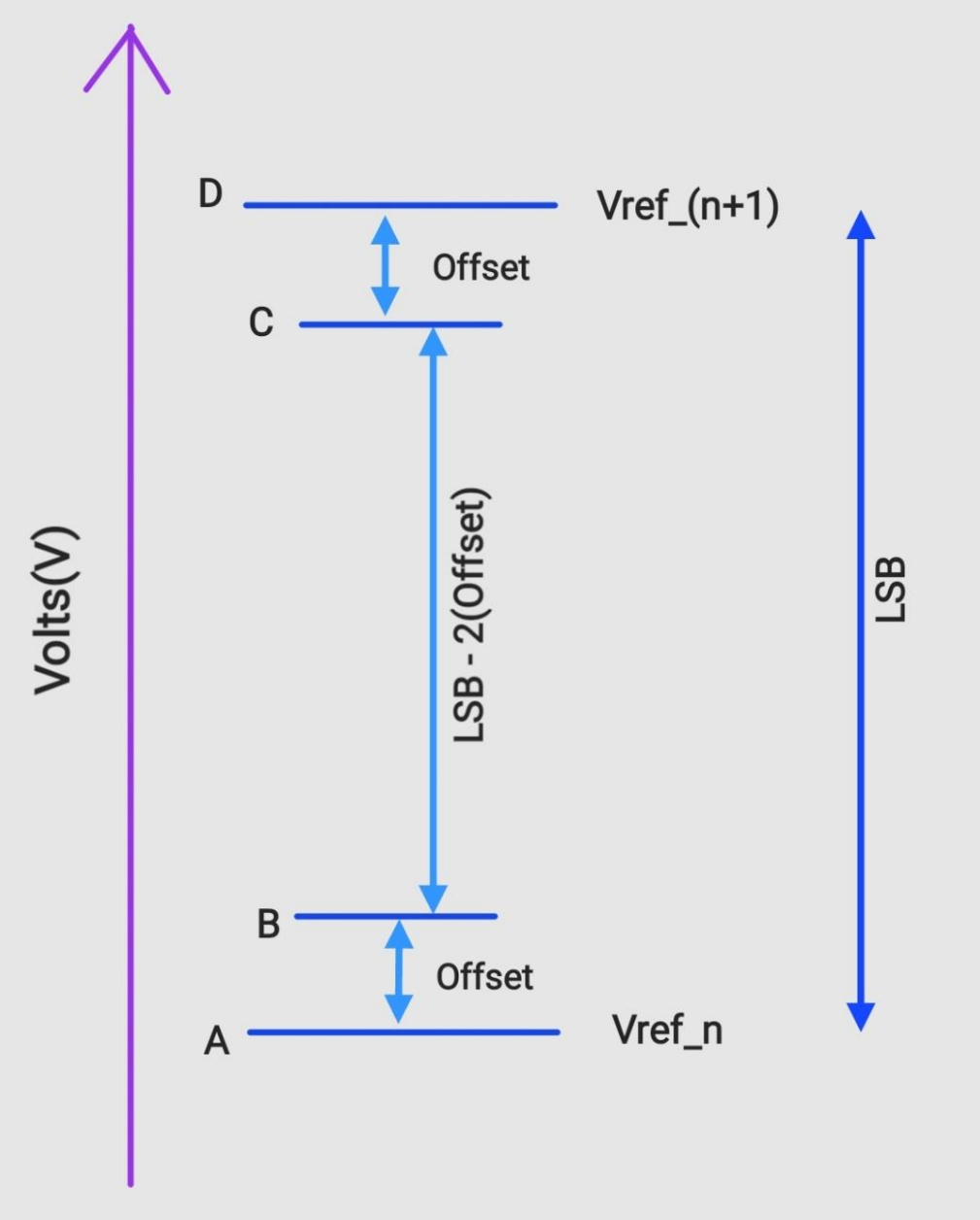
For the 15-SA array with V_{LSB} of 20mv:

Peak current consumption is 3.5mA during the AZ phase.

Average power consumption per conversion is 1.8mW.

Peak power consumption is 3.15mW during the AZ phase.

We estimate the ENOB or the statistical 'average' bit precision of the ADC using a simple model. We assume that whenever the input differential is within the bound of offsets (±3mv), the probability of an incorrect decision is $\frac{1}{2}$. Observe the diagram below.



The input voltage range AD spans one V_{LSB}. AB and CD are ranges where the input differential is within the bounds of offset. If the applied input lies in BC, the 4-bit output is correct. If the same input lies in AB or CD, the obtained output may be correct to 3 or 4 bits equiprobably.

Hence, the ENOB to first order, assuming the probability of incorrect decision is $\frac{1}{2}$ when the input differential is within the bounds of offset, from simple weighted averaging is $\frac{(V_{LSB}-2|Offset|)*N+(2*|Offset|)*(N-0.5)}{V_{LSB}}$ or, *3.85*, where N is the number of bits in ADC i.e., 4 and |Offset| is the maximum offset of 3mv from every SA of the ADC.

****