A 14-bit, 30 MS/s Pipeline ADC Model with 1.2 V Supply, >72 dB Dynamic Range, and >65 dB Peak SNDR in 45 nm CMOS

EE 240C Final Project Report, 12 May 2023

Averal Kandala, SID: 3031828138

1. Architecture Selection

1.1. ADC Type

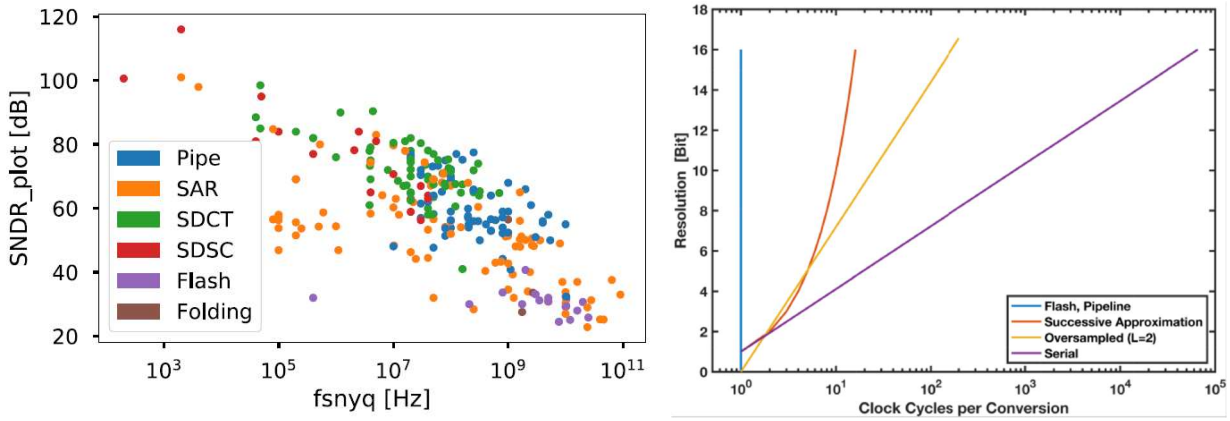

Fig. 1 | ADC performance comparisons, as presented in lecture [1,2].

The plots of Fig. 1, presented in lecture [1] and compiled in part from the data of [2], provide a succinct comparison of the capabilities of common ADC architectures as they pertain to the project. The project's bandwidth and SNDR specifications of ≥12.5 MHz and ≥65 dB, respectively, constrain satisfactory designs to the upper-right quadrant of the left plot. Notably, flash architectures, limited by comparator complexity that increases with resolution, and folding architectures, limited by nonidealities in the folding circuit, can achieve high speed but without sufficient resolution, marking them unsuitable for the project.

The remaining oversampled ("SDCT" and "SDSC" above), SAR, and pipeline ADC architectures offer the necessary performance with varying benefits and drawbacks. The right plot of Fig. 1 neatly illustrates how these different options trade off in speed and resolution. Oversampling inherently limits the Nyquist sampling rate to the circuit bandwidth divided by the oversampling rate, restricting application at higher input bandwidths. However, this strategy also reduces analog anti-aliasing filter requirements and allows exceedingly high resolution via noise shaping in the loop filter, with significant design complexity and susceptibility to fabrication errors being the main practical drawbacks. In contrast, the comparatively simple SAR ADC platform of binary search over the output of a capacitive DAC via a precise comparator can achieve relatively high resolution as well. The main limitations of the SAR approach are its "bit-at-a-time" conversion process, which constrains the sampling period to approximately the output word length times the clock period, and the accuracy of the DAC and comparator, which require special care in design, testing, and calibration. Finally, the pipeline ADC model splits the analog-

to-digital conversion into multiple stages, each containing a sub-ADC which digitizes its input for subtraction from the output residuum via a DAC. Although the latency introduced by the pipeline precludes this architecture's use in critical feedback applications, low-resolution flash sub-ADCs can be used to achieve high speed (throughput) and resolution without the complexity penalty of a full flash converter.

1.2. Pipeline Structure

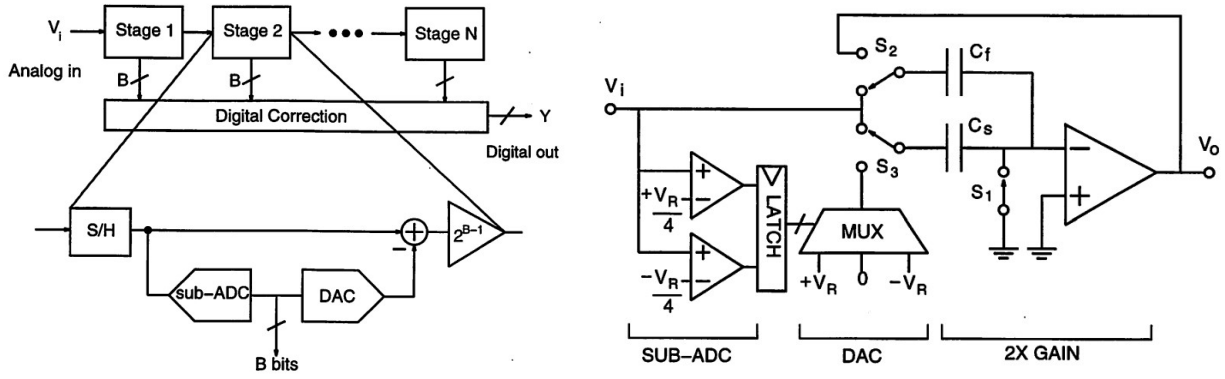

Fig. 2 | General pipeline model (left) and single-ended 1.5 bit/stage implementation (right) [3].

To investigate the limits of the versatile pipeline model, I created a design based on [3], which offers potentially maximal speed and simple digital bit correction, via bit shifts, through a cascade of instances of the robust, redundant "1.5 bit/stage" pipeline block shown above. Although the relatively low bandwidth specification could have allowed for an implementation with a higher effective number of bits/stage (each with higher gain and lower bandwidth) and that too with greater power efficiency, this optimum is shallow [4]. As a result, I opted in favor of the practicality and modularity of this design. Furthermore, no sample & hold amplifier (SHA) is included to minimize power consumption.

Each pipeline stage is implemented differentially to maximize dynamic range and common mode/noise tolerance, with differential modifications necessary to the capacitive DAC and switched-capacitor OTA gain stage depicted in the right plot of Fig. 2. The high stage redundancy enables the flash sub-ADC reference voltages to be "built into" the offsets of the (dynamic) comparators, potentially delivering significant power savings compared to a typical reference ladder [5]. The use of C_f $(= C_s)$ in both the sampling and charge redistribution phases increases the OTA loop feedback factor, β, to 1/2 in the ideal case, lowering the OTA noise contribution. Finally, a two-stage folded cascode topology (see Fig. 3) is selected for the OTA to ensure maximal input range and output swing, which sets V_{ref}. Note that the depicted circuit applies a different gain to the DAC output than the input ($\frac{C_s}{C_f}$, rather than $1 + \frac{C_s}{C_f}$), meaning that the DAC voltages must be adjusted accordingly in a model that applies the same gain to both signals.

2. Design Procedure and Analysis

2.1. Number of Bits

First, the dynamic range (DR) specification is used to establish the necessary number of bits in the ADC. By assuming that the thermal noise will be at the same level as the quantization noise (dealing a 3 dB hit to the SNR), we arrive at the following expression for the DR in terms of the minimum effective number of bits for the ADC ($ENOB_{min}$):

$$DR = 6.02 * ENOB_{min} + 1.76 - 3\ [dB] \quad (1)$$

For DR = 70 dB, this gives $ENOB_{min} = 11.83$ bits. To add leeway for distortion, which is not modeled in Eq. (1), and additional noise and nonidealities, I added approximately 2 bits to the $ENOB_{min}$ figure to yield 14 bits for the implemented converter, with 12 bits coming from the pipeline stages and 2 bits coming from a backend flash ADC.

2.2. Capacitor Scaling and Thermal Noise

As established in [4], to minimize system power, the capacitance should be scaled down by approximately the stage gain from one stage to the next, giving the relation in Eq. (3), where index $i = 0$ corresponds to the first stage.

$$G_{ideal} = 1 + \frac{C_s}{C_f} = 2 \quad (2)$$

$$C_{s,i} = C_{f,i} = \frac{C_{f,0}}{2^i}\ \forall i \in \{0, \dots, 12\} \quad (3)$$

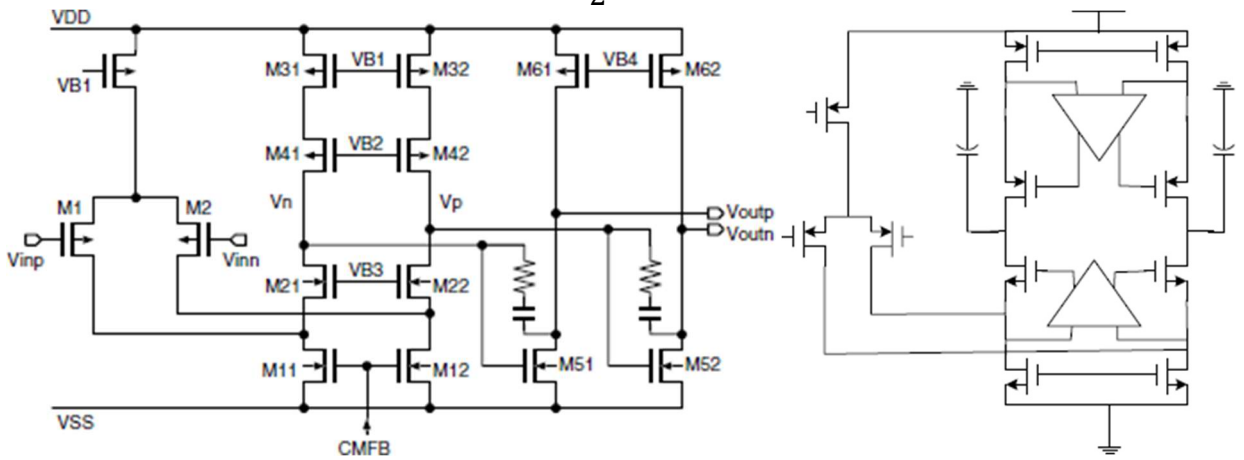

Fig. 3 | Modeled differential OTA, with a folded cascode input stage and common source output stage (left) [6]; gain-boosting (cascode device regulation) modification to the first stage (right) [7].

To derive the capacitor values for the first stage, and thereby those of all succeeding stages via Eq. (3), the first stage is arbitrarily selected to contribute NT_1 of the total input-referred thermal noise of the ADC, which had previously been assumed to equal the quantization noise.

$$\overline{v_{on,0}^2} = (G_{ideal})^2 NT_1 * N_{i,thermal} = (G_{ideal})^2 NT_1 * N_{i,quant} = \frac{(G_{ideal})^2 NT_1 \Delta^2}{12} = \frac{NT_1 \left(\frac{2V_{ref}}{2^{14}}\right)^2}{3} \quad (4)$$

The derived term in Eq. (4) is then equated to the analytical expression for the thermal noise at the output of the first stage, which is the sum of the sampling and OTA noise.

$$\overline{v_{on,0}^2} = \overline{v_{o,samp}^2} + \overline{v_{o,OTA,0}^2} = \frac{2kT}{C_{f,0}}\left(1 + \frac{C_{s,0}}{C_{f,0}}\right) + \frac{2\gamma N_1 kT}{\beta C_{c,0}} + \frac{2(\gamma N_2 + 1)kT}{C_{Ltot,0}} \quad (5)$$

$$\overline{v_{on,i}^2} = \overline{v_{o,OTA,i}^2} = \frac{2\gamma N_1 kT}{\beta C_{c,i}} + \frac{2(\gamma N_2 + 1)kT}{C_{Ltot,i}} \ \forall i \in \{1, \dots, 12\} \quad (6)$$

$$C_{Ltot,i} = 2C_{f,i+1} + (1-\beta)C_{f,i} \ \forall i \in \{0, \dots, 11\}, C_{Ltot,12} = 50\ fF \quad (7)$$

The first and second terms of the OTA noise in Eqs. (5) and (6) correspond to the first and second stages of the OTA, respectively. By fixing the second OTA stage's contribution NC_2 to 1/8 of the total OTA noise and using Eq. (7), we can solve for the feedback capacitor at the first pipeline stage:

$$C_{f,0} = \frac{1}{NC_2} \cdot \frac{2(\gamma N_2 + 1)kT}{(2-\beta)\overline{v_{on,0}^2}} \quad (8)$$

This then defines the rest of the system's capacitors through Eqs. (3), (5), and (6), with C_c referring to the OTA Miller capacitance. Note that no parasitic loading from the OTA input pair is considered in Eq. (7). However, in simulation, β was set to 1/3 rather than 1/2 to model this.

2.3. OTA Specification and Power Consumption

The two stage OTA depicted on the right of Fig. 3 allows a wide input range due to the folded cascode first stage and a wide output range due to the common source second stage, setting $V_{ref} = V_{DD} - 2V^*$. The G_m and unloaded gain, A_v, of this OTA can be approximated as follows, where g_m and r_o are assumed to be similar for all transistors in the signal path:

$$G_m = \frac{g_m(g_m r_o)^2}{4}; A_v = \frac{(g_m r_o)^3}{4} \quad (9)$$

With a g_m/I_d figure established through device characterization simulations and $N_{paths} = 4 + 2 = 6$ unit current paths, the total current in the *i*th OTA can be approximated as follows:

$$I_{OTA,i} = \frac{N_{paths} g_m}{\frac{g_m}{I_d}} = \frac{4N_{paths} G_{m,req,i}}{\left(\frac{g_m}{I_d}\right)(g_m r_o)^2} \ \forall i \in \{0, \dots, 12\} \quad (10)$$

With gain-boosting implemented through regulation of the cascode devices, each of G_m and A_v increases by a factor of ~$g_m r_o$ at the cost of 4 additional current paths. $G_{m,req,i}$ is set by a uniform GBW requirement in combination with the specific total load capacitance at the output node of each stage.

$$G_{m,req,i} = \frac{GBW_{req} C_{Ltot,i}}{\beta} \ \forall i \in \{0, \dots, 12\} \quad (11)$$

The required OTA GBW and gain are set by the dynamic and static settling errors, respectively.

$$\epsilon_s = \frac{1}{\beta A_{v,req}} \rightarrow A_{v,req} = \frac{1}{\beta \epsilon_s} \quad (12)$$

$$\epsilon_d = e^{-\frac{T_{clk}(2\pi GBW_{req})}{2}} \rightarrow GBW_{req} = -\frac{\ln(\epsilon_d)}{\pi T_{clk}} \quad (13)$$

3. Model Synthesis and Verification

3.1. Simulink and MATLAB Model

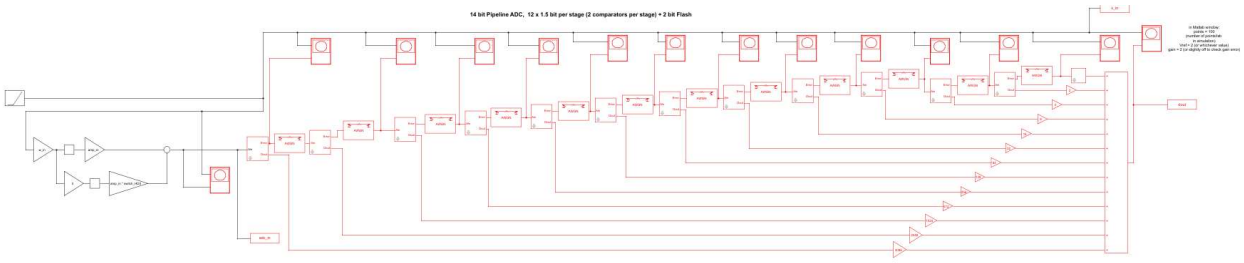

Fig. 4 | Simulink ADC implementation, simulation and parameter settings are controlled via MATLAB.

Using the Simulink pipeline stage and ADC blocks and MATLAB "glue" functions provided as part of the course [1], I implemented the proposed architecture in Simulink as shown in Fig. 4. For each of the 12 pipeline stages, the sub-ADC thresholds are $-\frac{V_{ref}}{4}$ and $\frac{V_{ref}}{4}$, while the DAC outputs are $-\frac{V_{ref}}{2}$, 0, and $\frac{V_{ref}}{2}$. The pipeline stage model applies the set gain to the input as well as the DAC output, which explains the discrepancy in DAC output settings compared to Fig. 2. The voltage thresholds of the 2-bit ideal flash ADC backend are $-\frac{V_{ref}}{2}$, 0, and $\frac{V_{ref}}{2}$. Digital stage outputs are bit shifted before being combined at the ADC output. The input is generated in discrete-time samples through the multiplication of each time point by the input frequency value followed by the sine operation, followed by a gain to set the amplitude. This allowed solver timestep issues to be avoided elegantly.

The equivalent output thermal noise of each pipeline stage, specified via variance (as calculated in Eqs. (5) and (6)), is introduced through AWGN (additive white Gaussian noise) blocks at each stage's error/residuum port. As mentioned previously, β is set to 1/3 rather than 1/2 to model the capacitive loading of the OTA input devices. Sub-ADCs and DACs are assumed to be perfectly ideal, with nonidealities in the gain stages realistically dominating. All necessary programmable parameters, including pipeline stage gain, noise variance, V_{ref}, and input sinusoid properties, are set via a controlling MATLAB test "harness" script which also runs simulations, collects output data, and performs spectral analysis. FFTs were taken with 2048 points, and the frequency of the input sine wave was pinned to (1021/2048) times the sampling frequency of 30 MHz, with data collected over 1021 cycles, to ensure the number of periods and the number of samples would be relatively prime, thereby preventing aliasing of harmonics.

3.2. Capacitor Mismatch and OTA Settling Error

For each stage, the gain is modeled in MATLAB as follows:

$$G_i = (G_{ideal} - \epsilon_c)\big(1 - (\epsilon_s + \epsilon_d)\big) \tag{14}$$

The right term contains the OTA settling error, which can only be negative. To accurately model the worst case gain, the capacitor mismatch term on the left is also taken to be negative accordingly. Using the findings of [8], the capacitor mismatch can be found from the unit capacitive variance.

$$\sigma_{cu} = 1\,\% \cdot fF \rightarrow \sigma_{\Delta c} = \sqrt{2}\sigma_{cu} \rightarrow \epsilon_{c,3\sigma} = 3\sqrt{2}\sigma_{cu} \quad (15)$$

Here, a 3σ tolerance has been allowed. Notice that this term becomes vanishingly small for even pF level capacitances. As a result, it was not found to significantly affect performance. In a system with higher capacitor mismatch, which typically does not vary with PVT, each stage's capacitive mismatch could be calibrated out through adjustments in the digital gain terms.

The OTA settling error terms set the required OTA GBW and gain as prescribed by Eqs. (12) and (13). In simulation, the minimal tolerable total OTA settling error was found to be 0.1%, and this error budget was evenly partitioned (as 0.05%) for static and dynamic settling to ensure that the required OTA GBW and gain were feasible.

3.3. Sampling Switch Nonlinearity

Sampling switch nonlinearity was implemented in the Simulink model through an additional path which triples the sine frequency before scaling by an HD3 term and summing with the ideal input. Ideally, the first two terms of the power series relation for a distorted differential signal would follow this form, with even harmonics cancelling out:

$$S_{od} = a_1 S_{id} + \frac{1}{4} a_3 S_{id}^3 \quad (16)$$

$$HD_3 = \frac{a_3 S_i^2}{4a_1} \quad (17)$$

However, when I attempted to implement this power series approach in the Simulink model, it resulted in a uniform boost of 3 dB to the SNDR, indicating that my implementation was incorrect. To ensure that some form of harmonic distortion was modeled in my system, I settled on the direct approach described above.

4. Simulation Results

4.1. Device Characterization

Based on Eqs. (12) and (13), as well as the OTA settling error budget described in the previous section, $A_{v,req} = 6000$ and $GBW_{req} = 72.6$ MHz. To ensure that these specifications are achievable in the chosen 45 nm process node, I created a single transistor testbench for the "g45n1lvt" and "g45p1lvt" devices in Cadence Virtuoso and used an OCEAN script (originally authored by Efthymios Papageorgiou) to sweep across different values of channel length, V_{gs}, V_{ds}, and V_{bs} in order to collect exhaustive device characterization data. I then plotted these data in MATLAB for $V_{ds} = 0.2\ V$ and $V_{bs} = 0\ V$ with the assistance of a MATLAB function authored by Efthymios Papageorgiou and analyzed trends in g_m, f_T, $g_m r_o$, and g_m/I_d for different channel lengths. As shown in Fig. 5, a channel length of 180 nm provided sufficient $g_m r_o$ and f_T to make the proposed design feasible, at least through the use of gain-boosting in the first OTA stage. In this analysis, $\frac{2\pi}{\tau} = \frac{f_T}{8}$ was used as a rough indicator for the GBW based on the discussion provided in lecture 9 [1].

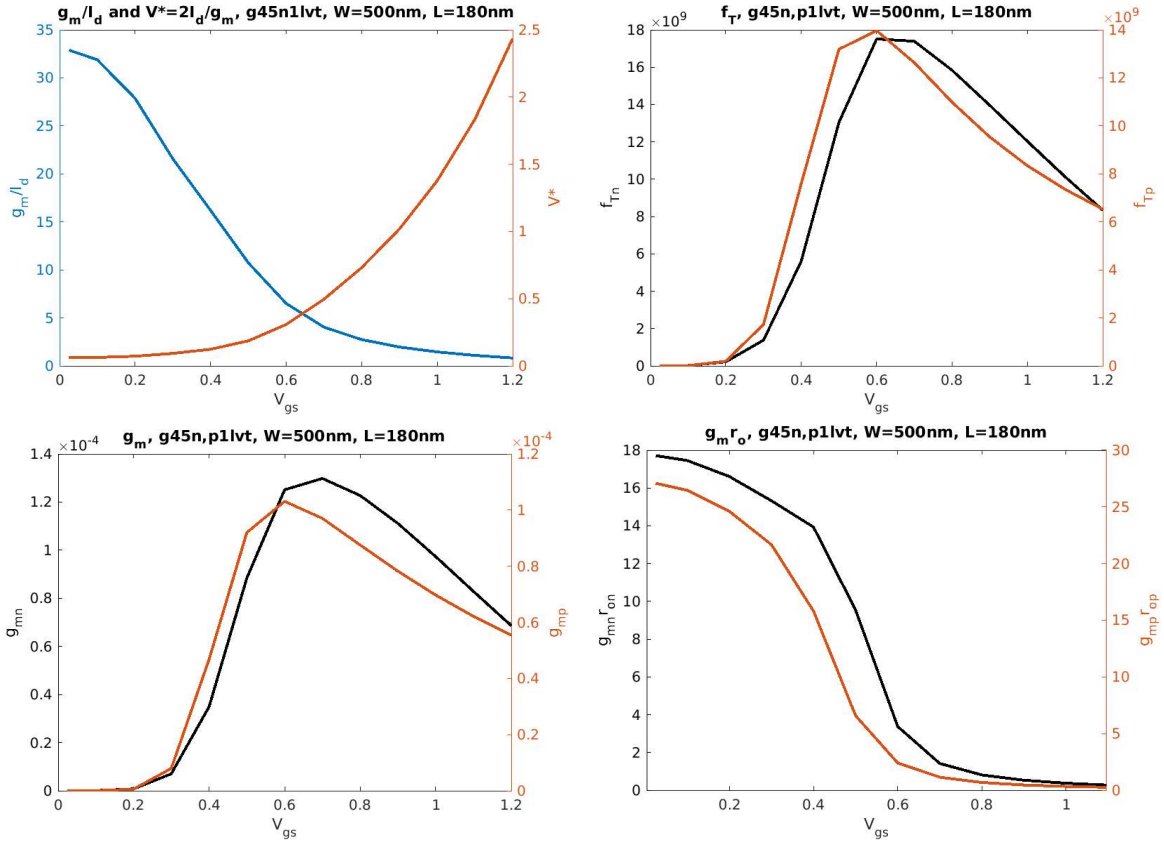

Fig. 5 | Biasing parameters for g45n1lvt and f_T, g_m, and $g_m r_o$ for g45n1lvt and g45p1lvt; plots and plotting scripts are original, but analysis scripts were originally authored by Efthymios Papageorgiou.

To reach V* = 0.1 V, as required to maximize output swing and place V_{ref} at 1 V, $V_{gs} = 0.4\,V$. At this bias, $g_m r_o$ is ~15, g_m/I_d is ~15 S/A, and f_T is ~8 GHz (note that these parameters remain relatively constant across device width). This intrinsic gain necessitates the implementation of gain-boosting, which then results in $A_v = 12656 > A_{v,req}$. The GBW requirement is met easily at around 1 GHz. Finally, modifying Eq. (10), we arrive at an updated OTA current equation:

$$I_{OTA,i} = \frac{4N_{paths}G_{m,req,i}}{\left(\frac{g_m}{I_d}\right)(g_m r_o)^3} = \frac{4 \cdot 10G_{m,req,i}}{\left(15\frac{S}{A}\right)(15)^3} = \left(8e - 4\frac{A}{S}\right) \cdot G_{m,req,i}\ \forall i \in \{0, \ldots, 12\} \qquad (18)$$

4.2. Sampling Switch Nonlinearity Simulation

To evaluate input sampling switch nonlinearity, Spectre testbenches for a CMOS switch and a bootstrapped NMOS switch ($V_{gs} = V_{DD}$) were prepared, with a sampling frequency of 25 MHz and an input frequency of (1021/2048)*(25 MHz), as well as a load capacitance of 411 pF. VDD was set to 1.2 V and the input was set to have an amplitude of 0.5 V on a 0.6 V (=VDD/2) common mode. Each device had a channel length of 45 nm, with NMOS widths nominally 5 microns and PMOS widths nominally 10 microns. Transient simulations were run over 1021 cycles for device (width) multipliers between 10 and 50, with data spectrally analyzed via FFT in MATLAB to extract HD3 values, presented in Fig. 6.

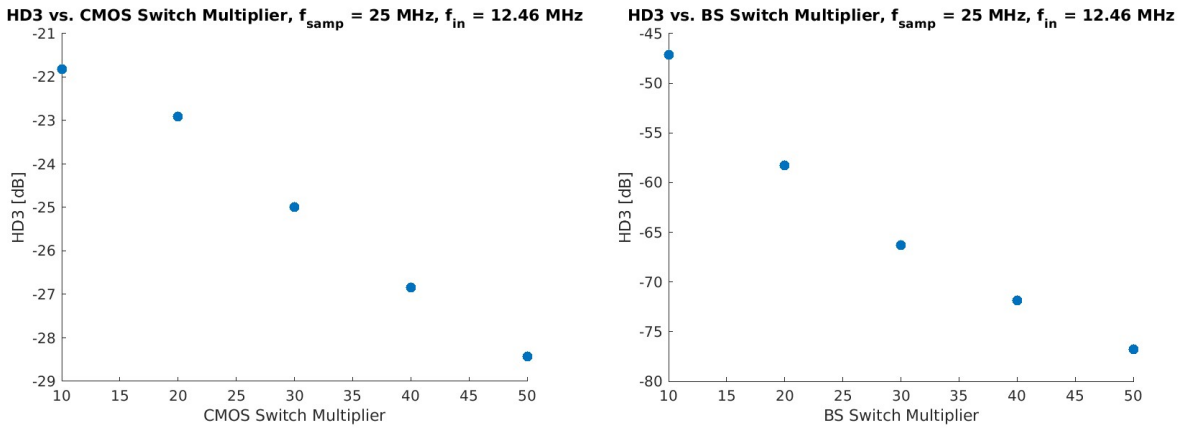

Fig. 5 | HD3 for a CMOS switch (left) compared to a bootstrapped NMOS switch (right).

The plots of Fig. 5 reveal that the two switch types are not even comparable for $f_s/f_{in}\sim 2$, and the bootstrapped switch is the far superior choice. The HD3 corresponding to a multiplier of 50 was used to add switch distortion in the Simulink model (1.4490e-4 in linear scale).

4.3. DFT Analysis and Dynamic Range Sweep

After confirming correct "staircase" quantization using a simple ramp input and establishing all nonideal parameters as described above, NT_1, the noise contribution of the first stage, was set to 1%, and the "harness" MATLAB script was run in a loop to sweep the input amplitude logarithmically (over 11 points up until $V_{ref} = 1\ V$) across the desired dynamic range. The plots, shown in Fig. 6, confirm that the project specifications are met by this design.

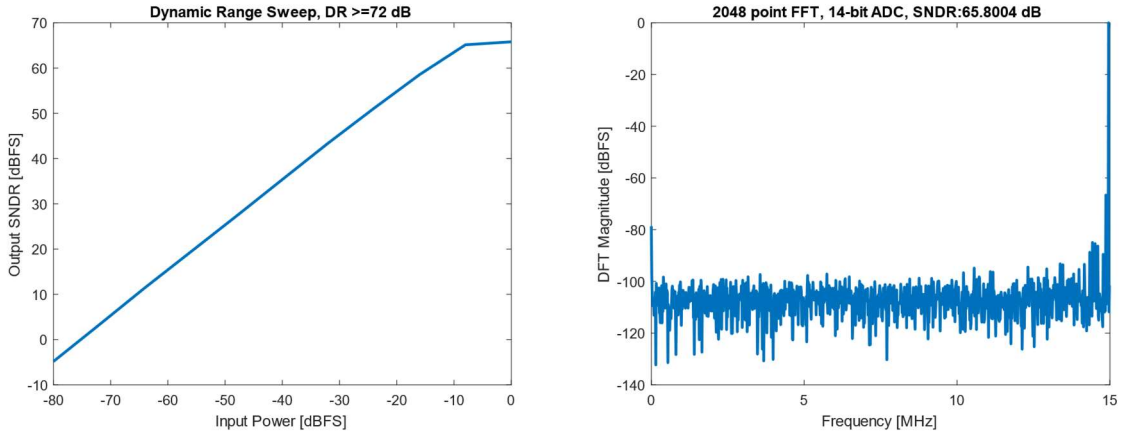

Fig. 6 | Dynamic range logarithmic sweep over 11 input powers (left) and FFT plot at peak SNDR (right) for the final design.

5. Performance Summary

5.1. Tables

Table 1: Capacitors and Power Consumption by Stage

Parameter	1	2	3	4	5	6	7	8	9	10	11	12
$C_f = C_s$ [pF]	2735	1368	684	342	171	85	43	21	11	5	3	1
C_c [pF]	2605*	1302	651	326	163	81	41	20	10	5	3	1
$G_{m,req}$ [mS]	992.6	496.3	248.2	124.1	62.0	31.0	15.5	7.8	3.9	1.9	1.0	0.2
I_{OTA} [μA]	794.1	397.1	198.5	99.3	49.6	24.8	12.4	6.2	3.1	1.6	0.8	0.2

Table 2: Performance Summary

Design Parameter	Specification	Actual
Bandwidth (BW)	≥12.5 MHz	≥15 MHz
Dynamic Range (DR)	≥70 dB	≥72 dB
Signal to Noise and Distortion Ratio (SNDR) at f_{in}~$f_s/2$	≥65 dB	≥65 dB
Nyquist Sampling Frequency	≥25 MHz	≥30 MHz
Supply Voltage (VDD)	1.2 V	1.2 V
Total OTA Current	Minimum	1.6 mA
Power Consumption (P)	Minimum	1.92 mW

5.2. Design Critique

This design process illustrated to me the danger of building an entire design on unchecked assumptions, as I realized on the last day of the term that I had not correctly accounted for the differential initial sampling noise in my calculations and needed to re-iterate on my design to meet the project specifications. This resulted in a few large parameter shifts, such as in the first stage noise contribution, NT_1, which shifted from 50% to 1%, as well as in the necessary capacitor values, which rose from the hundreds of pF to nF at maximum. In addition, the calculation for C_c in the first stage is likely incorrect, as it does not account for the sampling noise contribution to the output thermal noise of the first stage, hence the * in Table 1. These issues, along with the suspiciously low total current and power consumption, are likely indicative of flaws in my design reasoning and/or calculations. The value of the diverse perspectives that come as part of working in a team is quite clear.

From a technical perspective, many important behaviors in a pipeline architecture remained unmodeled in this implementation. Dynamic comparator power consumption and capacitive loading, as well as nonidealities in reference generation, are prime examples. Analyses of slewing limits and OTA compensation (which would directly constrain values of C_c) would be extremely important in a real design but could not be accommodated given the time limitations of the project. However, all this being said, this experience provides a useful starting point in ADC design, which I hope to reference as I continue in my career.

References

1. J. Vanderhaegen, EE 240C Spring 2023 Course Materials
2. B. Murmann, "ADC Performance Survey 1997-2023," [Online]. Available: https://github.com/bmurmann/ADC-survey
3. A. Abo, “Design for Reliability of Low-Voltage, Switched-Capacitor Circuits”, PhD Thesis, University of California, Berkeley, 1999. Available as UCB/ERL M99/25: http://www2.eecs.berkeley.edu/Pubs/TechRpts/1999/ERL-99-25.pdf
4. Yun Chiu, P. R. Gray and B. Nikolic, "A 14-b 12-MS/s CMOS pipeline ADC with over 100-dB SFDR," in IEEE Journal of Solid-State Circuits, vol. 39, no. 12, pp. 2139-2151, Dec. 2004, doi: 10.1109/JSSC.2004.836232.
5. T. Cho, “Low-Power Low-Voltage Analog-to-Digital Conversion Techniques Using Pipelined Architectures”, PhD Thesis, University of California, Berkeley, 1995. Available as UCB/ERL M95/23: http://www2.eecs.berkeley.edu/Pubs/TechRpts/1995/ERL-95-23.pdf
6. H. Ishii, K. Tanabe and T. Iida, "A 1.0 V 40mW 10b 100MS/s pipeline ADC in 90nm CMOS," Proceedings of the IEEE 2005 Custom Integrated Circuits Conference, 2005., San Jose, CA, USA, 2005, pp. 395-398, doi: 10.1109/CICC.2005.1568688.
7. Liu Zao, Cheng Jun, Zhang Hong and Chen Guican, "The design and optimization of gain-boosted OTA for high speed and high accuracy sample and hold amplifier," 2007 7th International Conference on ASIC, Guilin, 2007, pp. 461-464, doi: 10.1109/ICASIC.2007.4415667.
8. V. Tripathi and B. Murmann, "Mismatch Characterization of Small Metal Fringe Capacitors," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 61, no. 8, pp. 2236-2242, Aug. 2014, doi: 10.1109/TCSI.2014.2332264.