Low Quiescent Power CMOS Op-Amp in 0.5µm Technology

Kevin C. Fronczak

***Abstract*— This paper analyzes a low quiescent power CMOS operational amplifier design. The circuit architecture utilizes an inverting-style output to achieve 209 µW of static power consumption. A table of specifications is introduced and the performance of the amplifier over corners is presented. Due to the nature of the low power design, bandwidth was severely limited, but all other parameters fell very close to their target value. Various circuit improvements are presented as a way to increase the performance of this architecture.**

I. Introduction

Operational Amplifiers that operate with low quiescent power are becoming increasingly important in today's consumer electronics philosophy shift of "Always On, Always Connected"[1]. There are multiple novel ways to approach low quiescent power design, one of which is investigated thoroughly in this paper. The problem with any low power design is that some design specifications may need to be sacrificed to achieve the desired static power consumption.

The desired specifications for this opamp were an open-loop voltage gain of 3000 V/V, Phase Margin of greater than 50°, Unity Gain Bandwidth of 20 MHz, Voltage Swing of at least 500 mV away from the rails, Common-mode input range that includes either VDD or VSS, Quiescent Power Consumption of less than 200 µW, slew rate of at least 3 V/µs, PSRR at 60 Hz of at least -60 dB and PSRR at 1 MHz of at least -40 dB. The specification compliance table is introduced later in section IV, Table 1.

II. Circuit Architecture And Theory of Operation

The overall philosophy of the Op-Amp was to eliminate any unnecessary current branches. Figure 1 illustrates the architecture in block diagram form. A folded cascode architecture was chosen as the first stage due to the gain having an inverse relationship to current (Figure 2). Equation 1 shows this relationship.

$$(1)\ A_v = {g_m}^2 {r_{ds}}^2 = {g_m}^2 \frac{1}{\lambda^2 {I_d}^2}$$

This stage then feeds its single output into two common-source amplifiers with active PMOS loads. These two transistors help to increase the gain from the output of the folded cascode (but by a small amount in comparison). The main goal of these amplifiers was to essentially act as a level shift buffer before entering the output stage (Figure 3). The DC node voltage at the output of the folded cascode sits at roughly -1.56 V, measured experimentally. In a design where adaptive biasing techniques are used (see section VI. Discussion), this does not create a problem. However, the technique used in this particular design was to operate the output stage at subthreshold until a large input swing (caused by the large gain of the folded cascode) turns on either the NMOS or PMOS output transistors to source/sink the correct amount of current. Thus, if a -1.56 V node voltage sat on the PMOS output transistor gate, it would always be conducting, even in a quiescent state. If one were to level shift this voltage, it would also negatively affect the performance of the folded cascode since the wide-swing NMOS mirror would triode. The only solution, in this case, is to add some sort of buffer to allow for this -1.56 V node voltage while still holding the output transistors in a subthreshold region of operation.

The final crux of this architecture, DC bias networks aside, were MN17 and MP3 (shown in Figure 3). These diode-connected transistors act as pull-down/pull-up resistors, respectively, to help maintain a gate voltage on M26 and M27 that keeps them in weak inversion. As the voltage level at the gate of M27 rises, the voltage is pulled up by the PMOS load M24 which turns M27 on and allows it to sink current from the load. This voltage is the same on the gate of M26 and, conversely, as this gate voltage rises, M26 begins to turn off so that only M27 is sinking current. The opposite case occurs when the voltage swings low, M26 turns on and M27 turns off which means the PMOS output transistor, M26, sources current to the load. This essentially operates as an inverter since a low-swing on the 'input' causes a high-swing on the 'output' and vice-versa. Also, similar to CMOS logic, the only appreciable current draw is during voltage transitions which imply that the output stage is signal-dependant. This is a very useful feature for low quiescent power designs since it permits the designer to allow for large current *in transient conditions only*. However, if the 'input' of this inverter output stage does not have those pull-up and pull-down transistors, the voltage will sit at an undesired level determined by the V_{eff}

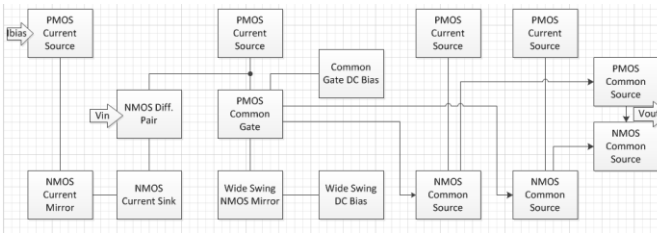

Figure 1. Op-amp Architecture Block Diagram

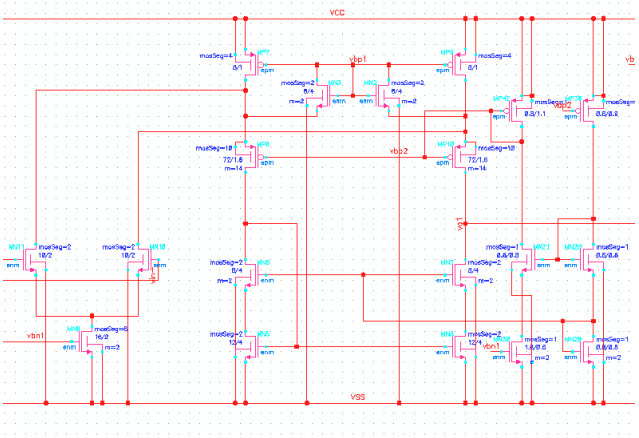

Figure 2. Folded Cascode stage with DC bias network

(and, by extension, the aspect ratios) of the common source and active load transistors. These level shift diode-connected transistors solve this issue by yanking the gate voltages either high or low when no signal swing occurs.

One concern, discussed more thoroughly in section IV. Specification Compliance, is that the high-impedance of the folded cascode output creates a very large dominant pole. Since this node is driving a common source amplifier, care must be taken to reduce the Miller Capacitance seen at that node. Two options are available: decrease R_{out} of the cascode stage (and decrease gain) or decrease g_m of the Common Source stage. Decreasing g_m presents multiple options (decreasing W/L, decrease I_D, etc) but in each case, gain decreases. Modifying the widths and lengths is the safest route to go as the aspect ratio can be maintained (to leave the gain unchanged), but the area can be modified which will decrease the Miller Capacitance as shown in Equation 2. The problem with this approach is that if the gain is too high in the Common-Source amplifier, the small modification in C_{gd} will be almost negligible.

$$(2)\quad C_{eq} \approx A_v C_{gd} = g_m\left(r_{dsn} || r_{dsp}\right) WLC_{ox}$$

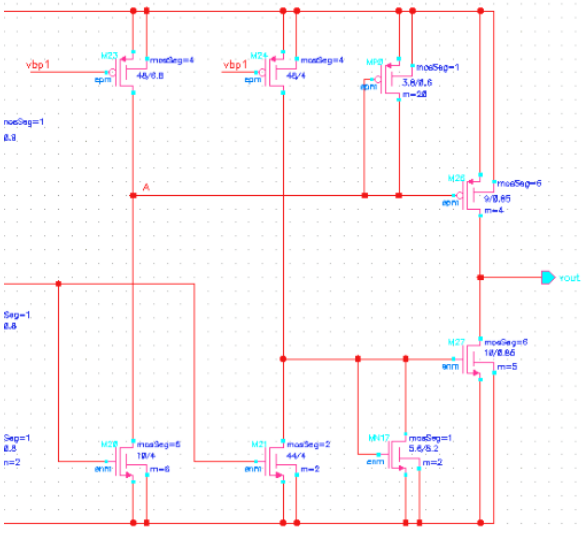

Figure 3. Common Source Level-Shift and Output Stage

III. Bias Networks

A large portion of this architecture was set aside for current and voltage bias purposes. Out of 34 total transistors, only 8 of them actually carry a signal. The remaining 26 either act as active loads or voltage bias generators.

The first major bias network is shown in Figure 4. Initially, the network only consisted of MP8, MP4 and MN4 in a simply current mirror configuration. The problem with that configuration was that since there were only two transistors between VDD and VSS, there was a large amount of current that had to be dropped across MP4 (and thus a large amount of current draw) in order to generate the correct bias voltage at the gate of MN4. The solution was to simply place a cascode current mirror at the drains of MP8 and MP4. This allowed for the current in that bias branch to be dropped from 8 µA down to about 3 µA while maintaining the same voltage levels at vbp1 and vbn1. The cascode mirror was chosen because no bias voltages needed to be generated (unlike a wide-swing cascode, for example) and that it required the total voltage drop to be at least $3V_{eff} + 3V_t$ below VDD by the time MN4 was reached. Given that the V_{eff} of all the transistors were roughly 0.25 V and the V_{tp} was around 0.9 V, the upper limit of vbn1 was -0.95 V and the lower end, given a V_{tn} of 0.7 V, was -1.55 V. The nominal voltage of vbn1 was -1.45 volts, so this range was tight enough to be acceptable.

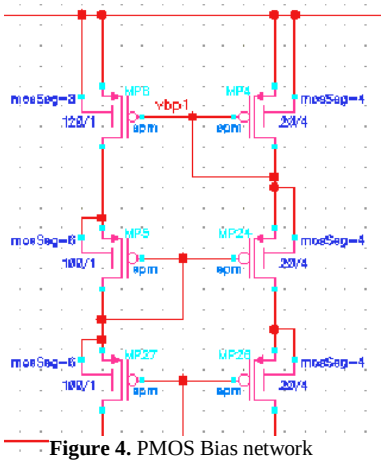

Figure 4. PMOS Bias network

The wide-swing mirror on the folded cascode stage needed to be biased externally, as did the common-gate pair. Originally, the biasing was achieved by two independent circuits – essentially just two current mirrors – but, just like in the previous bias network, the current draw was problematic. Another issue was that in certain corners, the common-gate bias voltage would drift too high, turning them off, and in other the same would happen to the wide-swing mirror. These two problems were solved concurrently by making each biasing branch dependant on the other. This network is shown in figure 5. Using the vbn1 node generated by the network in Figure 4, the first branch was biased with an NMOS sink. On the drain of this NMOS sink sit two transistors, M21 and M42. M21 is used to mirror the current over to the second branch while M42 is a diode connected transistor that creates the bias voltage for the common-gate pair. The second branch

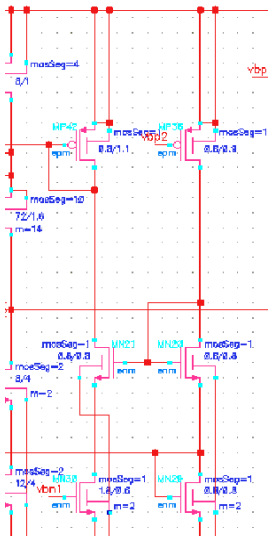

Figure 5. Folded Cascode Bias Network

has a PMOS current source that is biased from the same voltage node as the common-gate pair. Thus any voltage rises/drops will be matched in both branches and since the current is mirrored to a diode-connected transistor in the second branch, the currents stay the same as well. The final piece to this branch is MN29 which is simply a diode connected transistor that produces a bias voltage for the wide-swing mirror. Using this configuration allowed the current in each branch to be decreased from 6 µA to around 600 nA and helped to maintain a relatively constant *ratio* of bias voltages on the common-gate and wide-swing pairs.

IV. Simulation and Specification Compliance

Table 1 illustrates the desired specs, the worst-case achieved specs, and the corner at which that spec occurred. Any spec that failed to meet the target is highlighted in red.

As can be seen, only one spec fails by a significant margin and that was unity-gain bandwidth. This is caused by two issues, both of which were briefly mentioned in the Circuit Architecture section. The first being that as there is only 1µA of current flowing through the cascode transistors, the node resistance is incredibly high. Given a g_m of roughly 60 mS and r_{ds} of roughly 10 MΩ, the resistance seen at that node is about 6 GΩ. This large resistance, when coupled with the capacitance it drives, is the dominant pole of the system which will determine the -3dB frequency and, by extension, the unity-gain frequency. The capacitance that the folded cascode is driving is, using Equation 2, approximately 500 fF. This yields a dominant pole, f_{p1}, at roughly 50 Hz. In an absolutely ideal situation, no further pole would be encountered and this roll-off would yield a theoretical best-case unity-gain bandwidth of around 100 kHz. As Figure 6 shows, the -3dB frequency is closer to 100 Hz and the unity-gain ends up at 500 kHz. The discrepancies with the calculations are due to the current in the folded cascode actually being a bit higher than 1 µA. For example, if the current increased only to 1.2 µA, the pole would move to 70 Hz so a 20% increase in current results in a 40% increase in unity-gain bandwidth. However, as Table 1 illustrates, at a 1 µA nominal current, the power consumption is just slightly *over* spec so there really is no headroom to push the dominant pole outward.

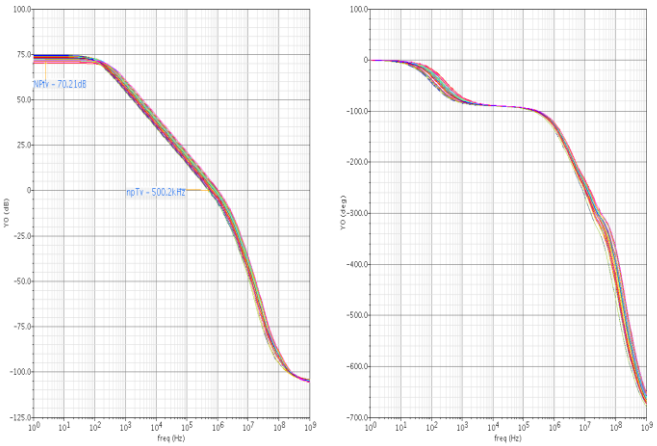

Figure 6. Bode Plot over corners

Param	Spec	Value	Corner
Gain	70 dB	70.2 dB	NPtv
UGB	20 MHz	500 kHz	npTv
PM	50°	55.3°	NPtV
SR	3 V/µs	2.27 V/µs (+) 3.00 V/µs (-)	npTv NPtV
Swing	500mV from VDD 500mV from VSS	314 mV 554 mV	NPTv abTV
PSRR	-60 dB at 60 Hz -40 dB at 1 MHz	-50.5 dB -40.35 dB	NPcV nPTV
CMIR	0-VDD or VSS-0	0 to VDD	-
Power	200 µW	209 µW	NPTV

Table 1. Specification Compliance Table

Figure 7 shows the Phase Margin plots over corners. The phase margin value is fairly consistent, but can be quickly improved with a compensation capacitor. However, this would decrease bandwidth and since that spec is already far below the desired and the Phase Margin meets spec, any compensation would end up hurting more than helping.

Figure 8 shows the Slew Rate plots. An obvious way to improve Slew would be to increase current in the folded cascode stage. This actually would have the added benefit of increasing the unity-gain bandwidth, as mentioned previously, while suffering just a small penalty to gain. Figure 9 illustrates the test-bench used to test slew rate. A square wave generator was attached to the positive terminal of the op-amp which was set-up in a unity-gain feedback configuration. The square-wave supplied a 1V pulse for 500 µs. The output voltage was then observed and slew-rate calculated by measuring the slope of both the rising and falling edges.

Figure 10 shows the large-signal voltage swing plots. As can be seen, the voltage rails at appropriate levels with only a few not making the negative swing spec. Since the worst only missed the specification by 50 mV, this was deemed to not be too much of a problem. The main contributor to the lack of swing in this architecture, primarily on the negative side, was the pull-down transistor. This is because it sets the gate voltage and, thus, the effective voltage for the output NMOS transistor. If sized to large, the NMOS doesn't turn on at the right time and the voltage swing never gets a chance to drop low enough. If sized to small, however, the NMOS will be released from weak inversion and will begin to conduct and sink an appreciable amount of current. Decreasing the size of the pull-down transistor will end up allowing the op-amp to meet the swing spec with only a very small hit to power consumption (if any at all).

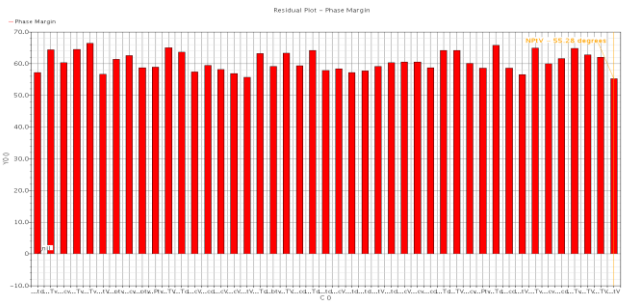

Figure 7. Phase Margin over corners

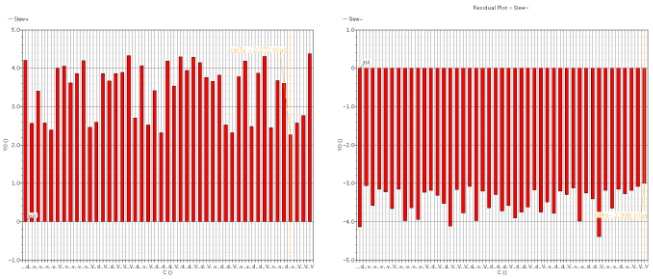

Figure 8. Slew Rate plots (rising edge, falling edge)

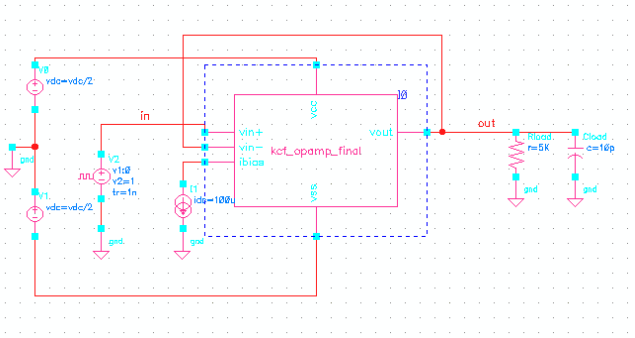

Figure 9. Slew Rate Test Bench

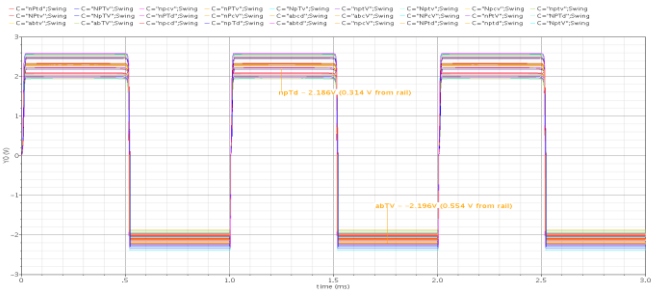

Figure 10. Large Signal Voltage Swing over Corners

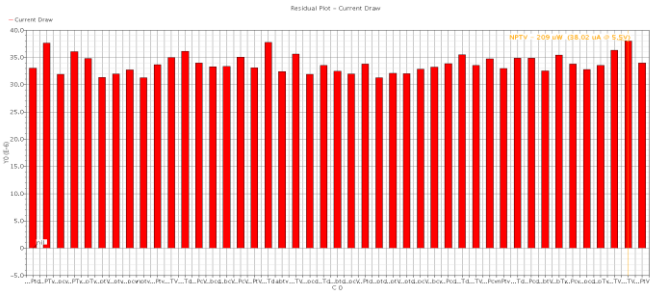

Figure 11. Quiescent Current Consumption over Corners

Figure 11 shows the quiescent current consumption (ignoring the biasing branch that sinks 100 uA externally). The current was measure simply by summing all of the node currents and plotting them for each corner. The total power consumption is then calculated on an individual basis by multiplying the total supply voltage by the current for that corner. Every case passed the 200 µW requirement except for high Temperature (125 ℃), High Supply Voltage (5.5 V) and hot process for both N and P. The test bench for the power consumption, large signal swing, gain, and phase is shown in Figure 12. The construction of the test bench is very straight-forward, a differential signal is applied between the positive and negative terminals of the op-amp and the output is measured for an AC sweep and in a transient case to determine the voltage swing.

Figures 13 and 14 show the PSRR for VDD and VSS while Figure 15 shows the test bench used for the VDD case. The rejection on the positive rail varied quite a bit in low frequency cases, but became much more consistent at high frequency. This is likely due to the various biasing schemes used that will change their node voltage slightly based due to variation in the supply. The noise is still rejected at a decent rate, -50 dB worst case, but could certainly be better. On the negative rail for 60 Hz, however, the rejection performed significantly better. This is likely due to the fact that there are simply more transistors this ripple needs to pass through before injecting into the signal path. At 1 MHz, the opamp performed quite well, meeting spec for both rails. A big contributor to this is likely the biasing network used for the cascode stage that essentially acts as a straight path to one rail or the other. This path helps to direct those high frequency ripples away from the sensitive signal path nodes. The test benches were very straight forward. A sinusoidal source was applied to either rail and then swept over frequency .The magnitude of the output voltage was then plotted against this frequency, as seen in Figures 13 and 14.

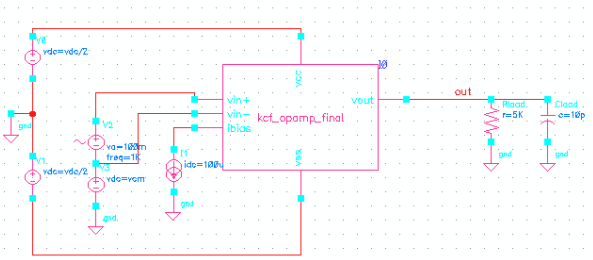

Figure 12. Open Loop Test Bench

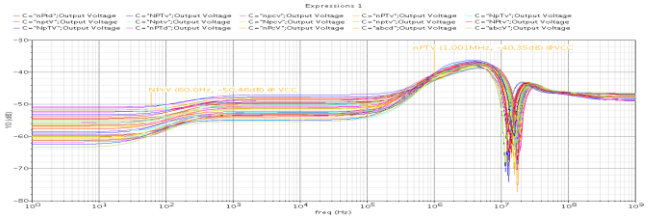

Figure 13. PSRR Over Corners for VDD

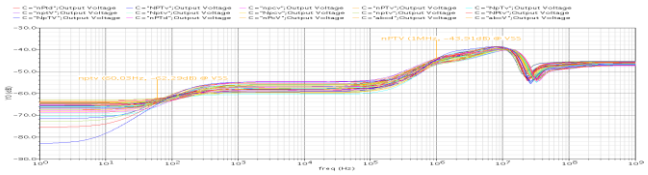

Figure 14. PSRR Over Corners for VSS

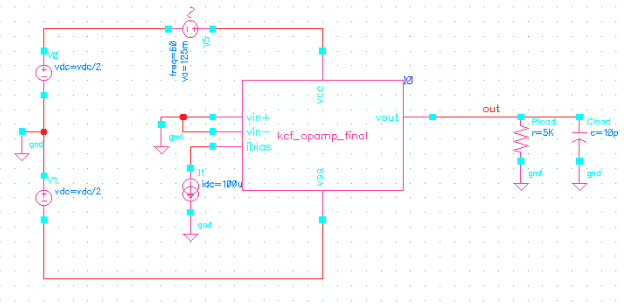

Figure 15. PSRR Test Bench

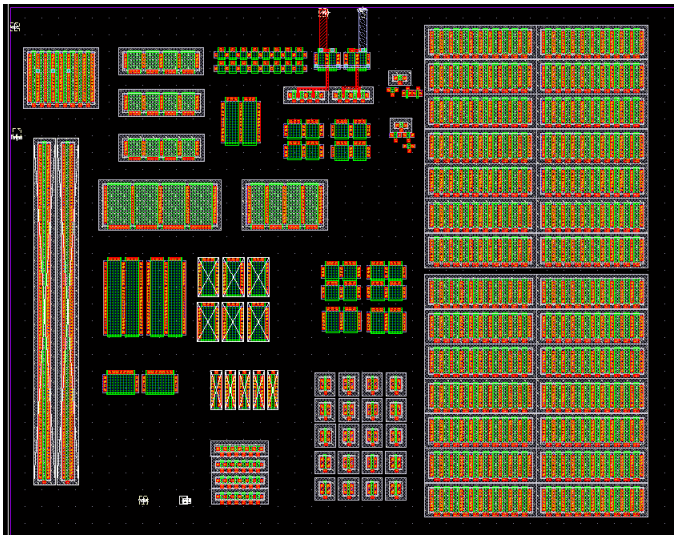

Figure 16. Physical Layout Floorplan

V. Physical Layout

Figure 16 shows the Physical Layout Floorplan. The input transistors are located at the top of the cell while the output is located at the bottom. The bias network is on the left edge and moves towards the middle of the package. The right-hand side of the cell is dedicated to the common-gate pair of the folded cascode stage. The area is roughly 4500 μm^2. Care was taken to minimize the size of the device as well as isolate the input node from the output node as much as possible.

VI. Discussion

In order to achieve most, if not all, of the specs, the best approach would be to adaptively bias the op-amp[2][3]. This allows for the output stage (or, theoretically, any stage) to be shut off when no input differential voltage is present. The basic idea is that the currents in each diff pair are mirrored and then subtracted through another transistor. This transistor will have no current when the differential currents match – which only happened in quiescent conditions. When a signal is applied, the voltages will differ and this difference in current will show up at that subtraction transistor. This current is then amplified and used to bias the output stage (or gain stage, etc). This essentially allows the Op-Amp to be designed without taking into account power-consumption as the bias network will simply turn everything off when no signal is applied. It's a very clever scheme that, if implemented properly, can work quite well. Based on the results in *A Novel Adaptive Biasing Scheme for CMOS Op-Amps*[2], the DC-gain was 90 dB, Phase Margin 85°, unity-gain Bandwidth 1.2 MHz and Power consumption of 24 µW. This scheme, if implemented exactly as the paper outlines, would not achieve all the desired specs outlined in Table 1 of this paper. However, it is not inconceivable that by increasing the current in a gain stage that the bandwidth could be pushed out with a hit to Phase Margin and Power Consumption. Given the fantastic values of both those parameters, this seems like a very logical path to take if implementing an adaptive-biasing scheme as an Op-Amp architecture.

VII. Conclusion

In conclusion, this paper presented a Low Quiescent Power architecture with 70 dB gain, 53° Phase Margin and 209 µW power consumption. Trying to achieve all desired specification was incredibly difficult for this specific architecture due to the limitations presented by a low power target. Bandwidth was the one specification that missed by a large margin, but cross-referencing to other low-power architectures, it appears that having a low GBP is not all that uncommon. Overall, this project was very successful in demonstrating the iterative design process inherent to analog integrated circuit design. It allowed for many different design paths which helped to solidify various topics and allowed for a much deeper and more fundamental understanding of MOSFETs in a transient setting.

VIII. References

[1] *Enabling Always On Always Connected Computing*, Kristoffer Fleming, Robert J. Hunter, Jon Inouye, Jeffery Schiffer, Intel 2002
[2] *A Novel Adaptive Biasing Scheme for CMOS Op-Amps*, Girish Kurkure and Aloke K. Dutta
[3] *A Low-Power adaptive biasing CMOS Operation Amplifier with enhanced DC-Gain*,Francesco Dalena, Vito Giannini, Andrea Baschirotto

TEST BENCHES:
Open-Loop-gain: /class/ee610/kcf2906/kcf_lib/bench_unity_gain
Slew_rate: /class/ee610/kcf2906/kcf_lib/bench_slew_rate
PSRR: /class/ee610/kcf2906/kcf_lib/bench_psrr

ADE_STATES:
Unity-gain: /class/ee610/kcf2906/ADE_states/BENCH/open_loop_gain
Slew-rate: /class/ee610/kcf2906/ADE_states/BENCH/slew_rate
PSRR: /class/ee610/kcf2906/ADE_states/BENCH/PSRR