EE6350: Class-D Audio Amplifier Report

Sarthak Kalani (sk3779)
Siddharth Shekar (ss4497)

May 29, 2014

This report discusses the transistor-level design and simulation results of a bridged three-level class-D amplifier. As opposed to conventional half-bridge and full-bridge variants of the class-D amplifier, this topology eliminates the need for the low pass filter at the output, thereby improving the efficiency of the amplifier. The top-level schematic for the system can be seen in Figure 1.

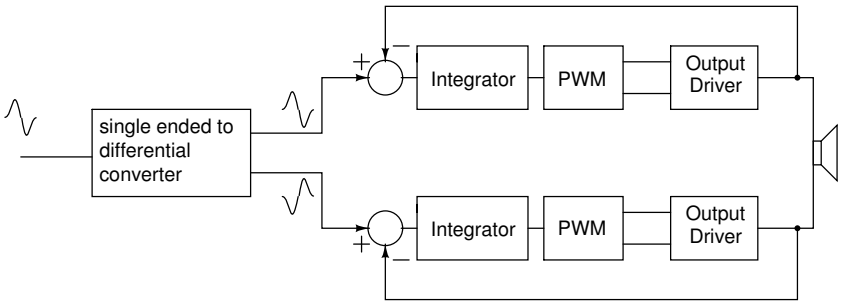

Figure 1: Top-level schematic of the class-D amplifier

1 System-level considerations

1.1 Input signal frequency range

Since the amplifier is to be used in the audio range, the input is expected to be between 20Hz and 20kHz.

1.2 Input signal amplitude

Since this is a closed loop design, the gain of the system is set by the feedback fraction. In this design, we chose the feedback fraction to be 0.5 so that the differential gain of the system is 2. The single-ended to differential converter offers a further "gain" of 2, so that the effective gain of the system from single-ended input to differential output is 4. Since the maximum output amplitude can be V_{DD}, the input amplitude needs to be limited to $V_{DD}/4$.

1.3 Integrator cut-off frequency

The integrator cut-off frequency needs to be low enough that it rejects spurious tones near the PWM frequency while simultaneously being high enough to let all of the input frequencies pass through unattenuated. In this design, this was set to be 80kHz.

1.4 Pulse Width Modulation frequency

The most important consideration while choosing the PWM frequency is the tradeoff between power dissipation and the suppression of high-frequency tones by the loop. As the frequency of the PWM increases, the integrator does a better job of suppressing the spurious tones but since digital power dissipation is proportional to fCV_{DD}^2, for a given capacitive load and supply voltage, the power dissipation goes up linearly with frequency. This reduces the efficiency of the system. On the other hand, if the PWM frequency is reduced, the efficiency improves, but now the integrator cannot reject the spurious tones near the PWM frequency as effectively.

The PWM frequency was chosen to be 500kHz. For an integrator cut-off frequency of 80kHz, this offers an attenuation of approximately 16dB at the PWM frequency. We note that in the bridged three-level topology, the nearest spurious tones appear at twice the PWM frequency instead of the PWM frequency itself. Therefore, the spurious tones see an effective 22dB of attenuation in the loop.

1.5 Stability analysis

If we model the loop assuming an ideal integrator with a UGB of $\omega_{UGB,int}$ and the PWM and output driver as a block that introduces a delay τ and a gain of G_{PWM}, the loop gain can be written as

$$L(s) = \frac{G_{PWM} e^{-j\omega\tau} \omega_{UGB,int}}{s}$$

In our implementation, $G_{PWM} = \frac{V_{DD}}{V_\Delta}$, $\omega_{UGB,int} = \frac{1}{R_F C}$. Therefore, the loop gain becomes

$$L(s) = \frac{V_{DD} e^{-j\omega\tau}}{j\omega V_\Delta R_F C}$$

The loop gain goes to 1 at a frequency of $\frac{V_{DD}}{V_\Delta R_F C}$. At this frequency, the phase is given by $-\frac{\pi}{2} - \frac{V_{DD}}{V_\Delta R_F C}\tau$. Setting $V_{DD} = 1.8V$, $V_\Delta = 0.9V$, $R_F = 200k\Omega$ and $C = 20pF$, for a 60° phase margin, we can tolerate a maximum delay of $\tau_{max} = 1\mu s$.

Note that this analysis assumes that the frequency where the loop gain goes to 1 is much smaller than the PWM frequency. In our case, the assumption holds true because the PWM frequency is 500kHz and the loop gain goes to 1 at approximately 80kHz. This assumption lets us write the PWM gain as $\frac{V_{DD}}{V_\Delta}$. This analysis also assumes that the OTA acts as an ideal op-amp. This will be largely true if the UGB of the OTA is much larger than the frequencies under consideration.

1.6 Noise and distortion

Due to the presence of the integrator in the loop, the noise contribution of the loop will be dominated primarily by the integrating OTA. Since the OTA's that are part of the single-ended to differential converter appear directly in the signal path, they will also directly add to the noise. However, given that we expect the signals to be large in amplitude, distortion effects will be stronger than noise.

The Total Harmonic Distortion (THD) of the amplifier is primarily affected by 2 things:

Dead-time The dead-time is defined as the period during which neither the pull-up nor the pull-down part of the output driver is active. Theoretically, if we had ideal switches, then the dead-time could have been made arbitrarily small. However, given that MOS transistors don't turn off until their V_{GS} drops below V_T and that the control signals itself will have a finite slope, we need to introduce some dead time intentionally so as to avoid turning on both pull-up and pull-down networks simultaneously which would effectively short V_{DD} to V_{SS}. This would result in a large transient current spike since these devices have very high current carrying capacity, and this might destroy the chip. Further, the amount of dead-time introduced needs to be enough so as to be able to avoid shoot-through currents even in the presence of PVT variations.

Triangle-wave linearity Given that in our implementation, the triangle wave is generated using an integrator, the slope of the triangle wave should be largely constant near its zero-crossings. Due to the finite bandwidth of the triangle wave generator, the peaks of the triangle wave will not be perfectly sharp. However, this will affect performance only if the signal is made large enough that it is compared to the peak of the triangle wave. This degradation in performance can be avoided by reducing the signal amplitude

2 Transistor-level design

2.1 Output Transconductance Amplifier

The OTA was designed first, since it is the block that is used in the most number of places. In order to allow for a large input amplitude swing, a folded cascode topology was chosen for the first stage. This also offers the advantage of high gain in the first stage. The second stage was chosen to be a common source amplifier with active load. The schematic for the 2-stage OTA is shown in Figure 2.

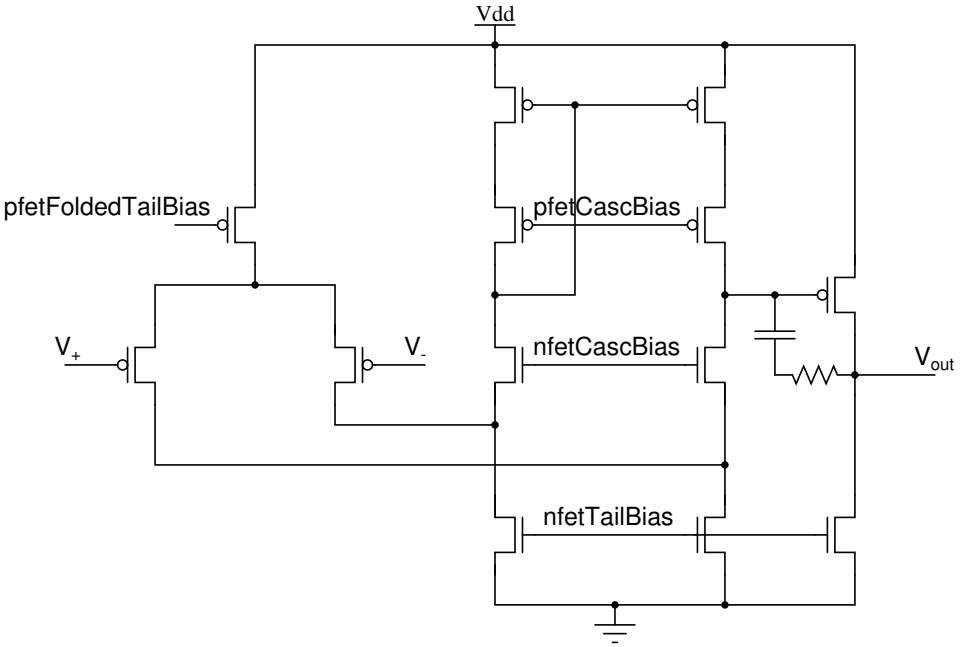

Figure 2: Schematic of the 2-stage OTA (bias circuits not shown)

The tail transistors were biased with an overdrive of 200mV to improve matching while the input transistors were biased in weak inversion so as to maximize their $\frac{g_m}{I_d}$. A 560fF Miller capacitor (VN-

CAP) was used to stabilize the OTA. Further, a 750Ω zero-cancelling resistor was added to improve the phase margin of the OTA.

2.2 Comparator

The comparator has 3 stages - a fast low-gain stage, a slow high-gain stage and an output stage with relatively high current drive capability. The schematic for the comparator is shown in Figure 3. Although the third stage is not needed if the second stage were equipped with high current drive capability, for the best matching, we would need to have the same-sized devices in the mirroring branch too. Hence a third stage was added to minimize the overall area of the comparator.

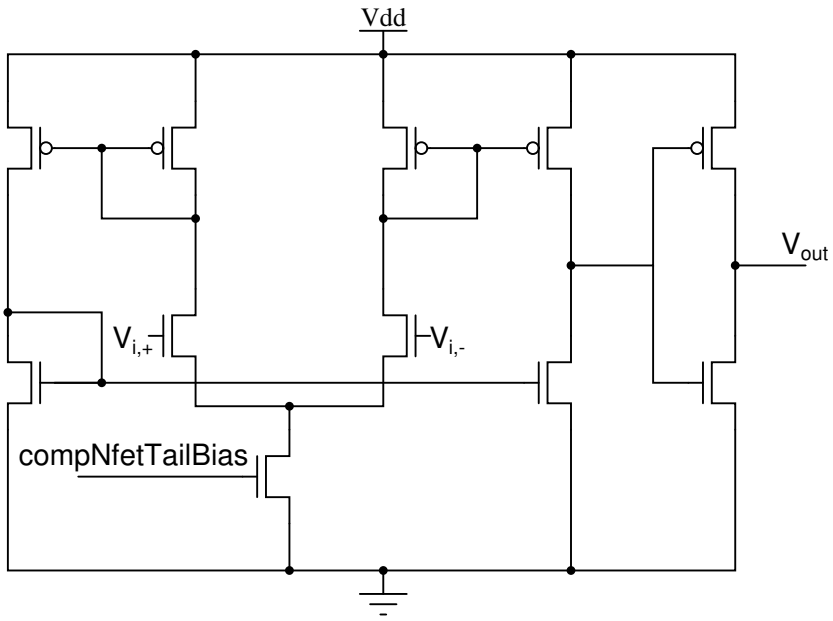

Figure 3: Schematic of the comparator

2.3 Bias circuit

A single bias circuit generates all the required potentials for the OTA and comparator circuits. Note that we decided to distribute voltages across the chip instead of currents. Cascode biases were generated by connecting 4 transistors in series to replicate the effect of $W/4L$. Tail current biases were generated by having a dummy cascode transistor (appropriately biased) so that the tail bias generating transistors see the same V_{DS} as the tail transistors in the actual circuits. A single 10μA current source needs to be connected to pfetFoldedTailBias to generate all the voltages. The schematic of the bias circuit is shown in Figure 4.

2.4 Single-ended to differential converter

The conversion from single-ended signal to differential signal was achieved using 2 OTAs as shown in Figure 5. R in the schematic was chosen to be approximately 10kΩ.

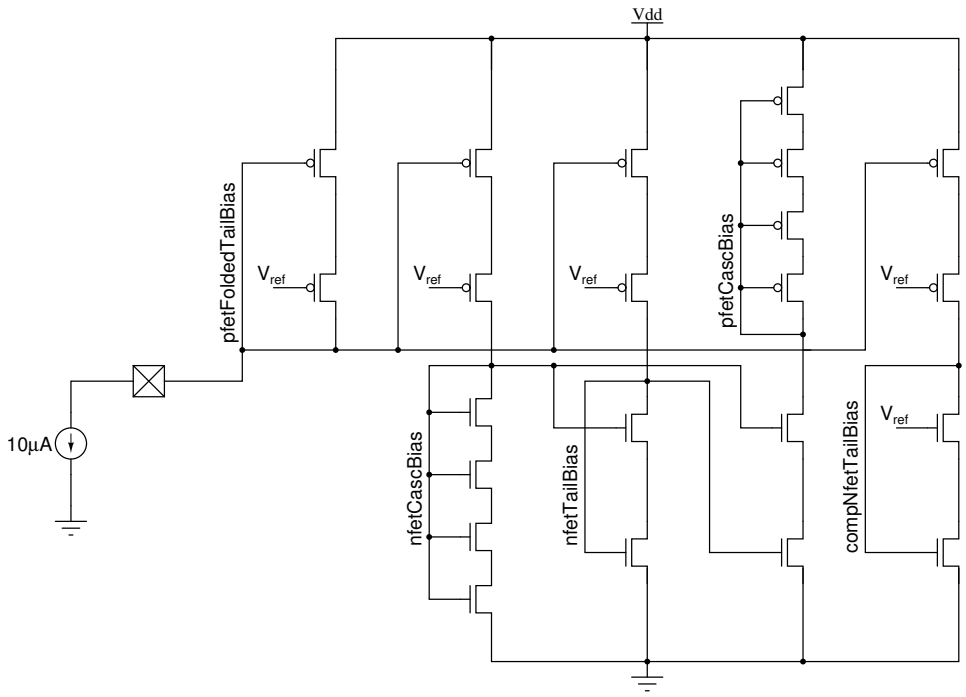

Figure 4: Schematic of the bias circuit

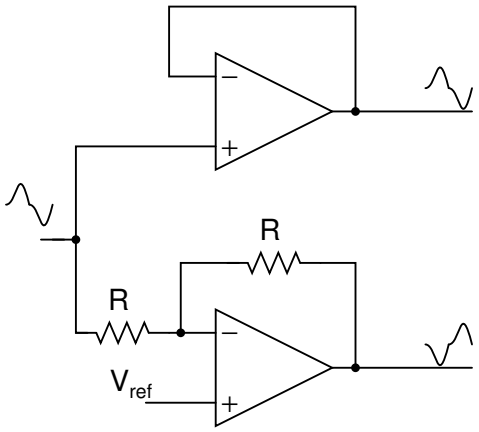

Figure 5: Schematic for the single-ended to differential converter

2.5 Triangle wave generator

The triangle wave generator is required in order to realize the pulse width modulation operation. It uses an integrator and a Schmitt trigger as shown in Figure 6. In this case, the peak-to-peak amplitude of the triangle wave becomes V_{DD}/a where 'a' is set by the resistor ratio shown in the schematic. Also, the time-period of the triangle wave is determined by R_1 and C and is given by $4R_1C/a$. In our implementation, we chose $R_1 \approx 100k\Omega$, $R_2 \approx 10k\Omega$, $a = 2$ and $C \approx 10pF$ giving a frequency close to 500kHz.

2.6 Pulse Width Modulator, Non-overlap Generator and Buffers

The schematic of the pulse width modulator, the non-overlap generator and the buffers are seen in Figure 7. The comparator shown in the schematic is the same as is described in Section 2.2. The digital gates are all standard CMOS structures with $1.8\mu m$ nFETs and $7.2\mu m$ pFETs and minimum length sized

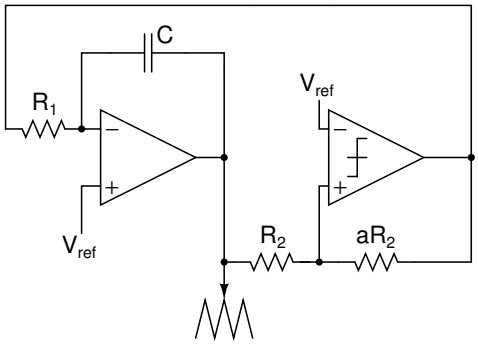

Figure 6: Schematic of the triangle wave generator

appropriately depending on the transistor stack height. $C = 20pF$ is a MIM-capacitor that is responsible for providing the 20ns non-overlap period that we need. Since the waveform at the output of the NAND gate is very slow, we use buffers to improve the "sharpness" of this waveform.

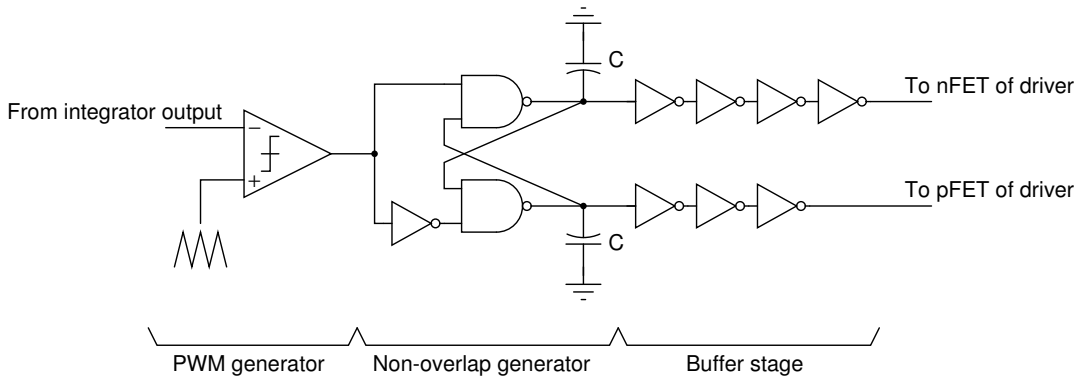

Figure 7: Circuit schematic of the PWM, non-overlap generator and buffers

Initially, the design was to operate with dual power supplies - 1.8V for the circuits up until the non-overlap generator and the subsequent digital circuits operating on 3.3V so that a larger amount of power could be delivered to the load (level shifter not shown in the schematic). We had sized the buffers so that they offered the same delay from the outputs of the non-overlap generator to the input of the respective transistors that they were driving. However, since the output drivers started switching too fast (in the order of 100's of ps), the current through the load would switch from 0 to 400mA ($3.3V/8\Omega \approx 400mA$) and considering a bondwire inductance of 6nH, this would lead to an $L\frac{di}{dt} = 6 \times 10^{-9} \times \frac{400 \times 10^{-3}}{400 \times 10^{-12}} = 6V$! As a result, on the chip, both V_{DD} and V_{SS} lines would see large transient voltage spikes when the output switched which led to breakdown of other analog transistors. This, coupled with fabrication problems, motivated us to change to a fully 1.8V design.

Due to the switch to a fully 1.8V design, the $L\frac{di}{dt}$ problem was mitigated but not as much as was required. In order to reduce its impact further, the buffers which were initially sized in a gradually increasing size to minimize delay were redesigned so that the output drivers would switch an order of magnitude slower. We note that after the redesign, the overall delay from the output of the non-overlap generator to the gates of the output drivers was mainly dominated by the delay across the last NOT gate part of the buffer chain. This delay was made the same for both the p and n paths so that the delay for either path remained approximately the same.

2.7 Output Drivers

The schematic of the output driver is shown in Figure 8. The primary consideration while determining the size of these transistors is the desired power efficiency. Let the drop across the load be denoted by V_{Load}. Then, the power drawn from the supply is $\frac{V_{DD}V_{Load}}{R_{Load}}$ while the power delivered to the load is $\frac{V_{Load}^2}{R_{Load}}$. The efficiency is then given by $\frac{V_{Load}}{V_{DD}}$. We can see that $V_{DD} - V_{Load}$ is the V_{DS} across the pFET output driver and the nFET output driver of the differential path. Therefore, for a given current through the load, the transistors were sized so that a minimum of 90% efficiency is achieved. This stage was designed with a 3.3V supply in mind and then converted without changing the dimensions when the decision to switch to 1.8V was made (to remove the need to re-layout this stage).

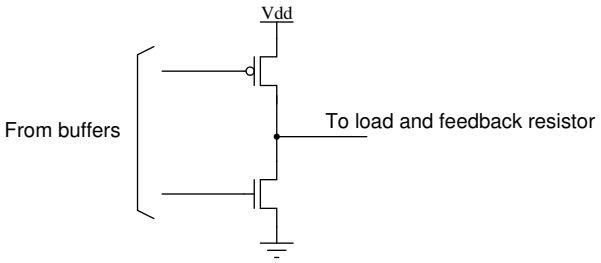

Figure 8: Schematic of the output driver

3 Layout

This section discusses the specific design choices made during the layout of each block. Layout image was followed, for the most part, for the lower metal layers with M1, M3 running vertically and PC, M2 running horizontally. M4 was primarily used for routing wires from pads to the circuit. MT and AM were respectively used for V_{SS} and V_{DD} respectively.

Since the design is inherently differential, the top-level layout reflects this. One half of the circuit was initially laid out and then was copied and mirrored in order to get the differential half. Common blocks like the bias circuit and triangle wave generator are placed along the line of symmetry. The top-level layout is seen in Figure 9. The line of differential symmetry is vertical and cuts across the middle of the chip. The signal enters from the bottom side of the layout and exits from the top. A "+" marker at the top-right is used to indicate the location of pin 1 and the required orientation with respect to the package.

5 RC clamps are placed at the vertices of the power rings with 2 clamps at the top-left vertex and 1 each at each of the other 3 vertices. ESD protection for all pins includes only HBM protection since adding the CDM resistor and diodes would have required more area. Further, it was not feasible for the output pins because it would introduce a series resistance which would have severely degraded performance.

3.1 OTA

The layout for the OTA block is shown in Figure 10. The transistors in the layout were arranged in roughly the same manner as the transistors in the schematic. The input signals of the OTA enter from the left and the output is taken from the right. Supply and ground buses run horizontally along the top and the bottom of the layout. The input pair was interdigitated and dummy devices were added at the edges so as to improve matching. The Miller capacitor was realized using a MIM capacitor. Also, the zero-cancelling resistor was split into 4 resistors of the same value with 2 each in series forming 2R and

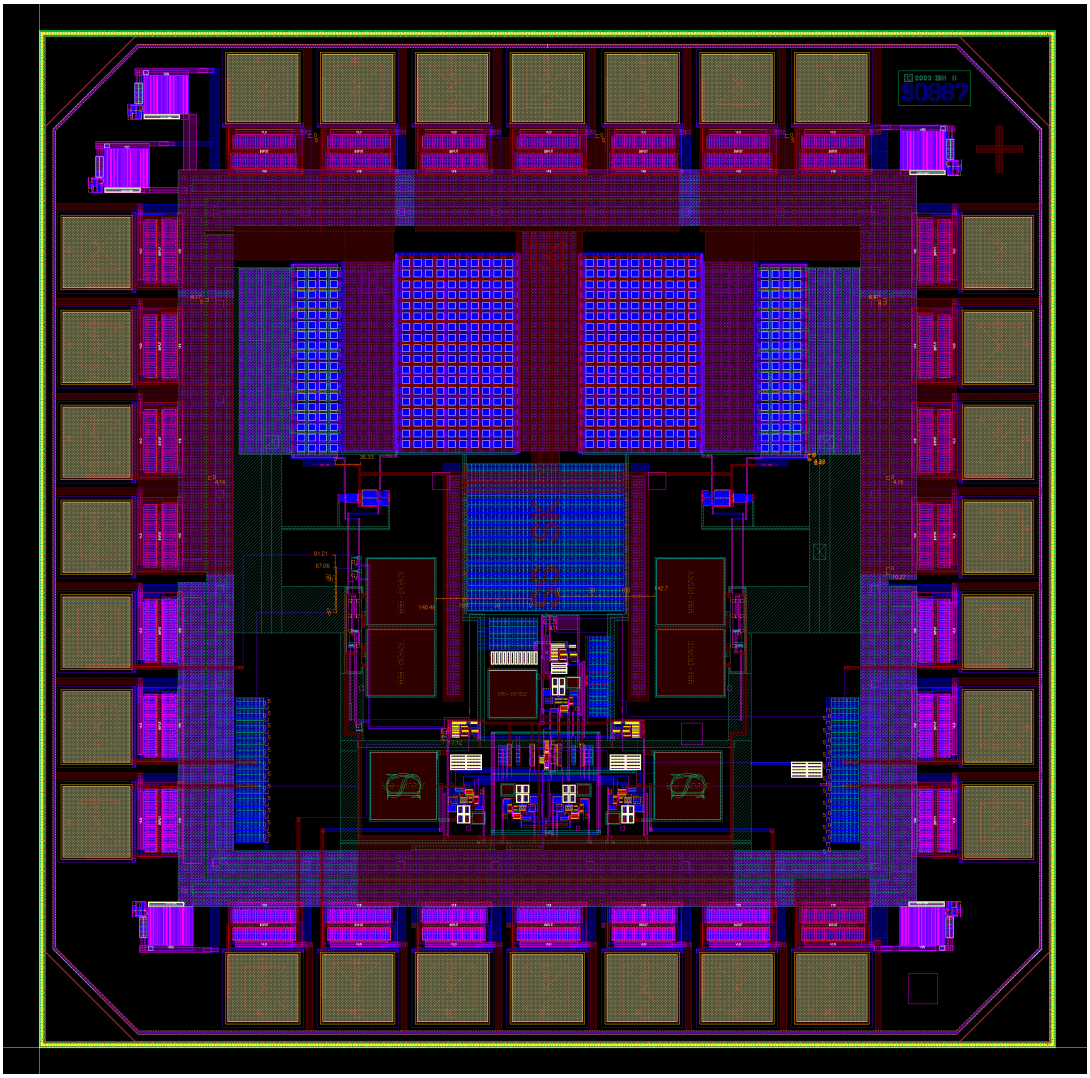

Figure 9: Top-level layout

the two 2R's in parallel forming the desired R. This arrangement improves the accuracy of the value as opposed to just using a single resistor of value R.

3.2 Comparator

The layout of the comparator block can be seen in Figure 11. As was the case with the OTA, the layout transistors were arranged in largely the same way as the transistors in the schematic. The inputs of the comparator come in from the left and the output leaves from the right. The V_{DD} and V_{SS} buses run horizontally along the top and bottom of the layout respectively. Several substrate contacts were added so that the substrate has a good connection to V_{SS}.

3.3 Bias circuit

The layout of the bias circuit is shown in Figure 12.

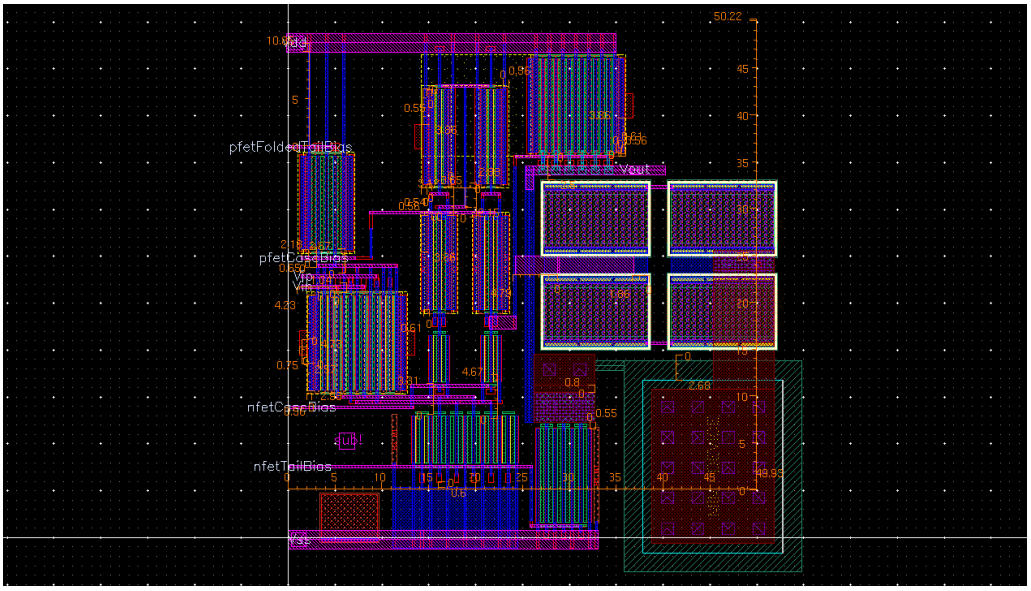

Figure 10: Layout of the OTA block

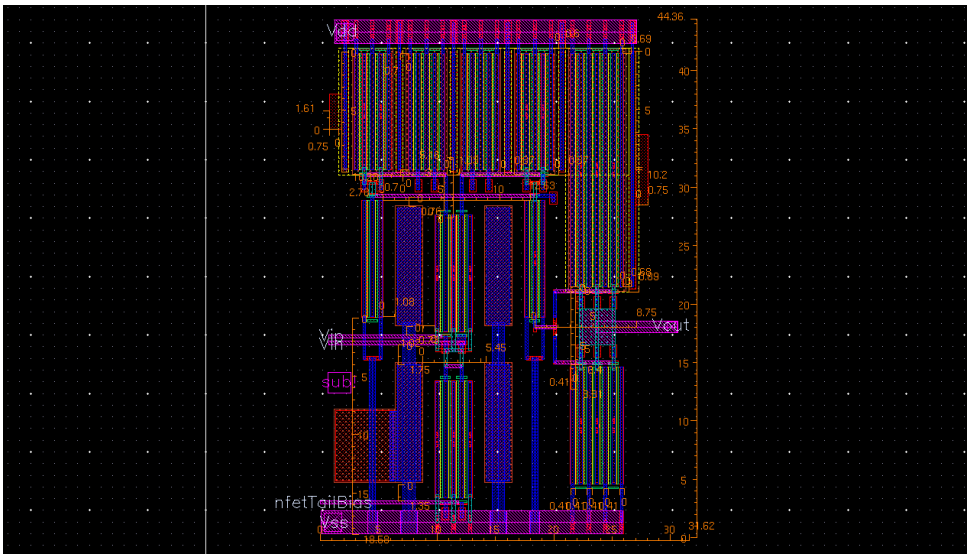

Figure 11: Layout of the comparator block

3.4 Single-ended to differential converter

The layout of the single-ended to differential converter is shown in Figure 13. This layout was created by mirroring the OTA layout about the X-axis. Care was taken to maintain the differential nature of the design. The V_{DD} and V_{SS} buses run vertically along the left and are tapped off as required to power the OTA blocks.

3.5 Triangle wave generator

The layout of the triangle wave generator is shown in Figure 14. The resistors 'aR_2' and 'R_2' (refer Figure 6) are placed close to each other to improve matching. The large resistor R_1 is broken up into 12 pieces and meandered to improve accuracy of the realized value.

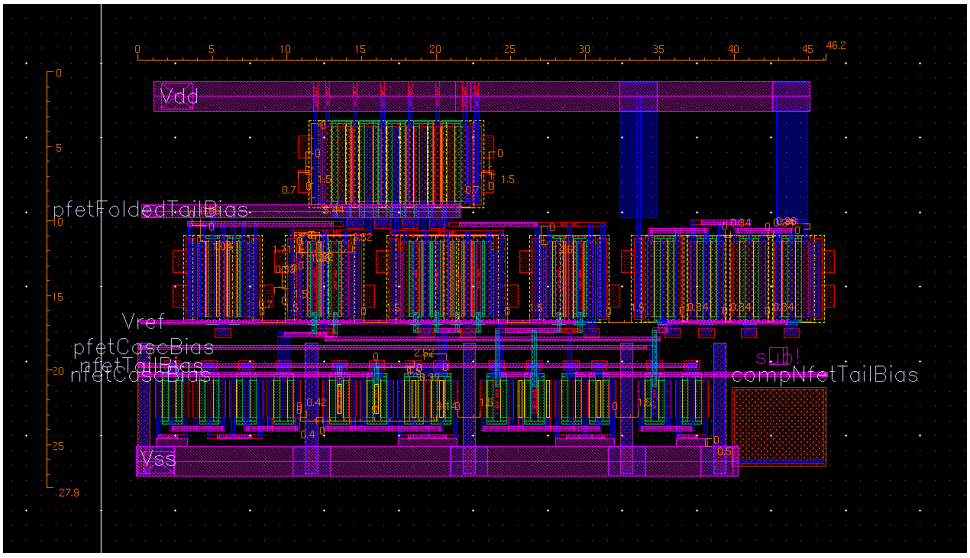

Figure 12: Layout of the bias circuit block

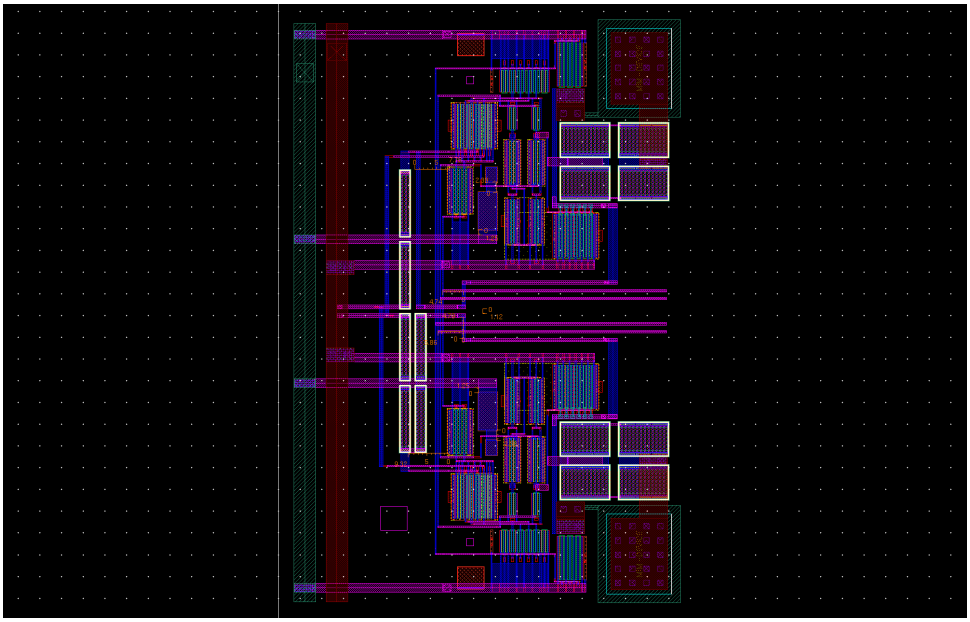

Figure 13: Layout of the single-ended to differential converter

3.6 Non-overlap generator

The layout of the non-overlap generator can be seen in Figure 15. The layout area is mostly dominated by the MIM capacitors that provide the required delay. The relatively tiny digital gates are all located at the top of the layout in the same order as they appear in the schematic (Figure 7).

3.7 Output drivers

The layout of the output drivers is shown in Figure 16. Starting from the top, the layout consists of the V_{SS} bus, the nMOS pull-down, the output bus, the pFET pull-up and the V_{DD} bus. Metal layers are strapped together through the use of via arrays so as to minimize the wiring resistance and to increase the current-carrying capacity of the bus.

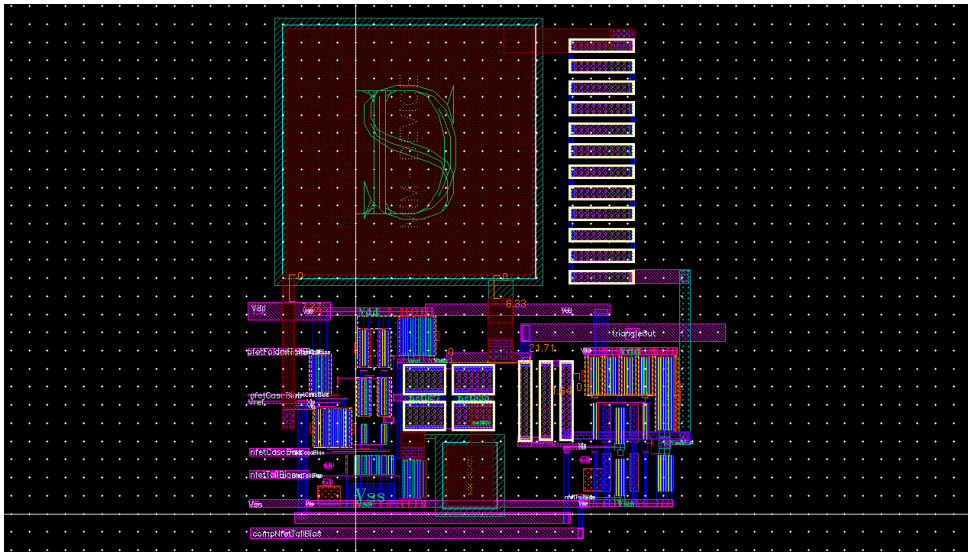

Figure 14: Layout of the triangle wave generator block

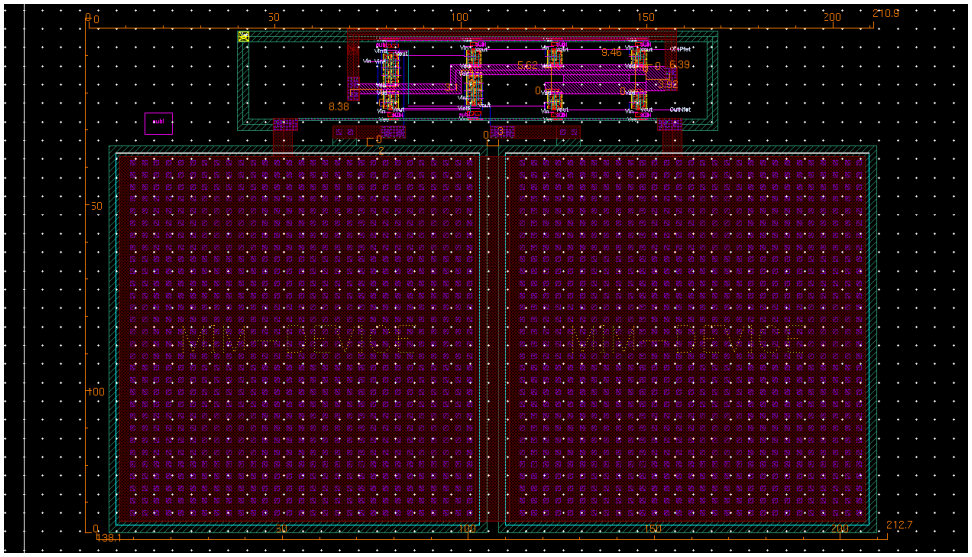

Figure 15: Layout of the non-overlap generator block

Each of the nFET pull-down and pFET pull-up transistors are laid out as an array of multi-finger transistors. Alternating supply and output buses form a grid through the transistor array in order to efficiently distribute power and minimize resistance. Care was taken to make the inner wires wide enough so as to withstand the relatively large currents that they would have to carry. M4 was used exclusively for the output bus with MT and AM being dedicated for the V_{SS} and V_{DD} buses respectively.

4 Simulation Results

4.1 OTA

Figure 17 shows that the open-loop phase margin for the OTA is approximately 83° with a UGB of 77MHz. The OTA consumes 300μA of current from a 1.8V supply resulting in a power dissipation of 540μW.

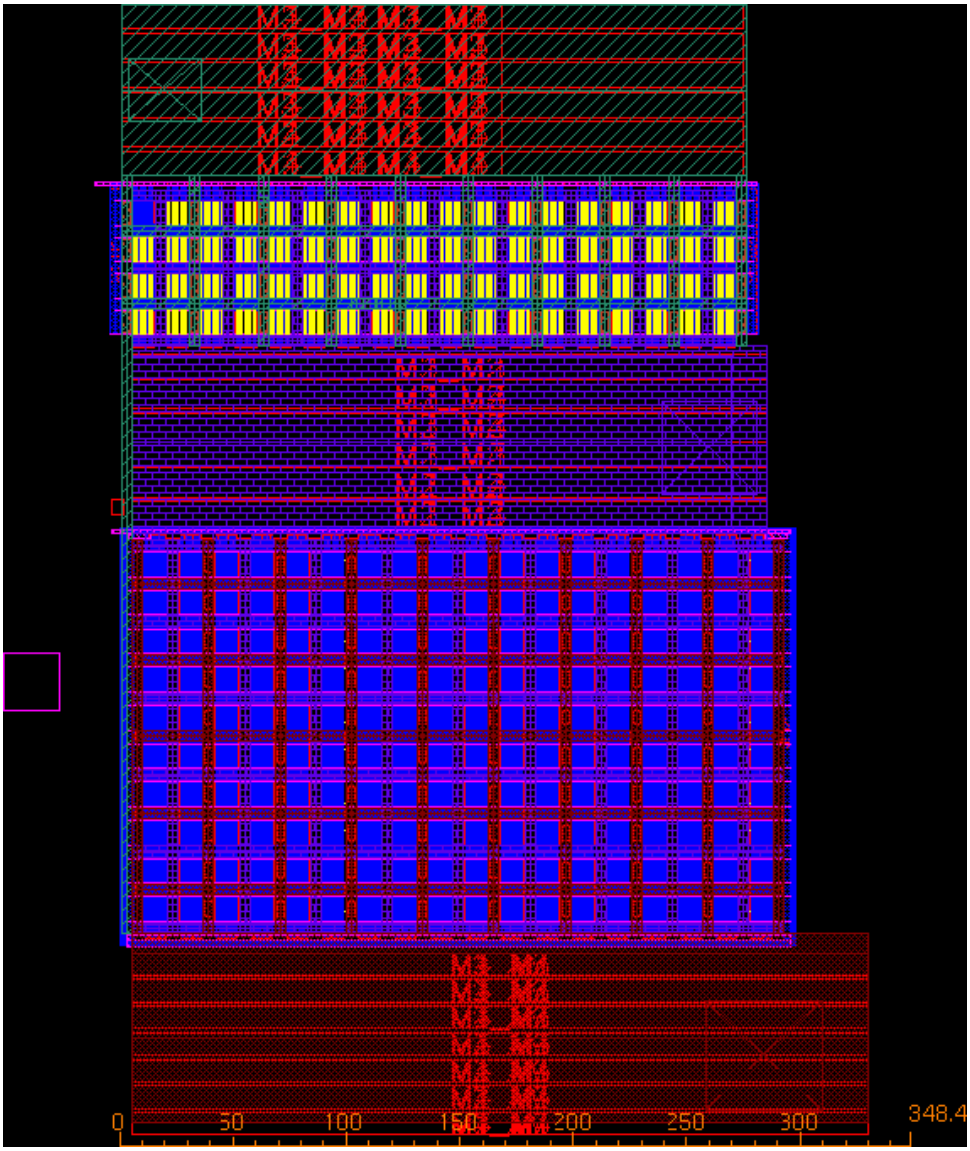

Figure 16: Layout of the output driver stage

4.2 Comparator

Figure 18 shows the transient response of the comparator with a typical $20k\Omega$ load. The rise time is 930ps with a delay of 1.4ns and the fall time is 2.56ns with a delay of 6.1ns.

4.3 Single-ended to differential converter

Figure 19 shows the transient response of the single-ended to differential converter block. As expected, the block generates an output that is in phase with the input and an output that is 180° out of phase with the input.

4.4 Triangle wave generator

Figure 20 shows the transient output of the triangle wave generator block. The block generates a triangle wave with a peak-to-peak amplitude of $V_{DD}/2 = 0.9$V and a frequency of approximately 500kHz.

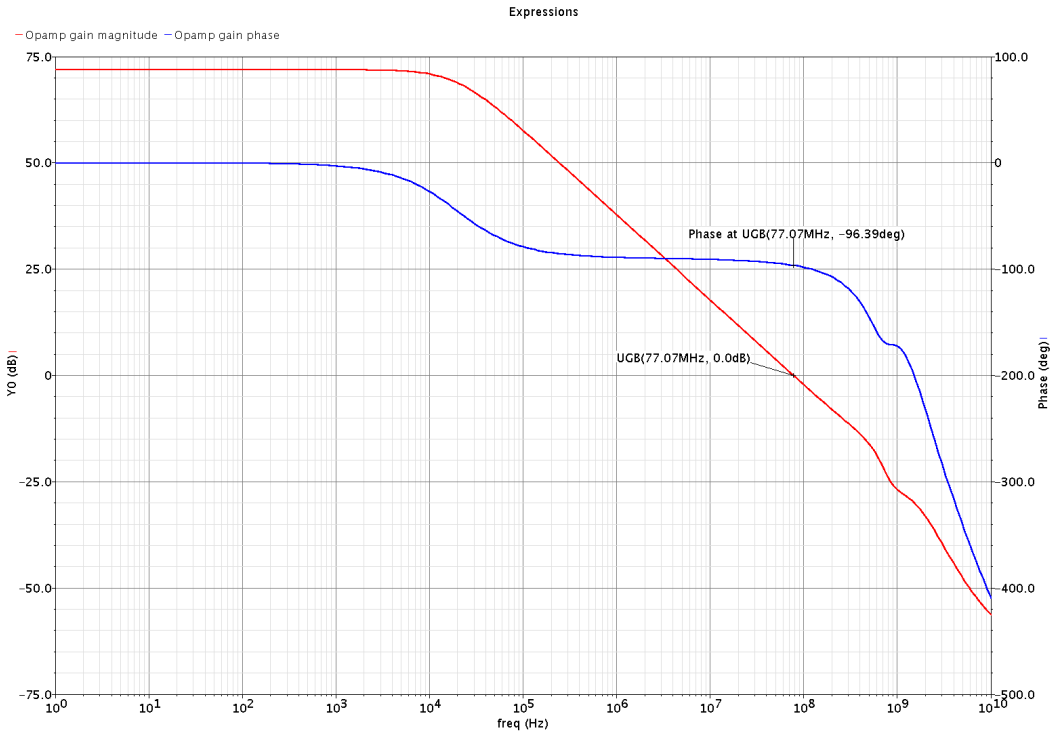

Figure 17: Plot showing the open-loop phase margin and UGB of the OTA

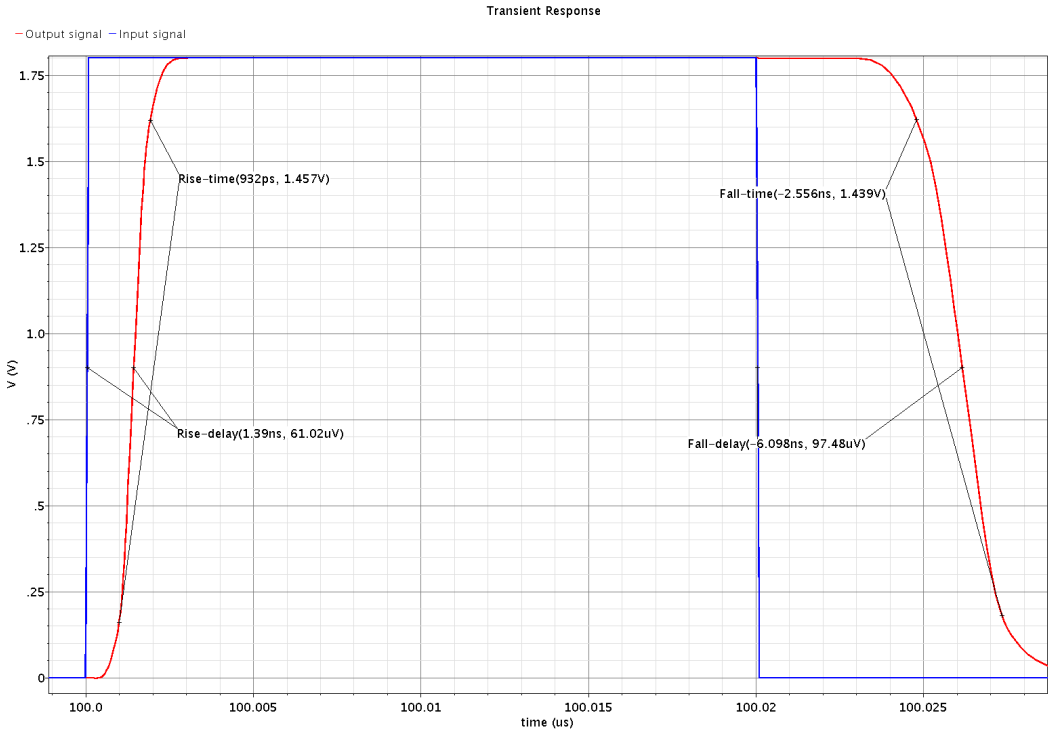

Figure 18: Plot showing the transient response of the comparator with a $20k\Omega$ load

4.5 Non-overlap generator

Figure 21 plots the input and output voltages for the non-overlap generator block (including buffers) in the SF corner (which was deemed to be the worst-case). The block adds a non-overlap period of

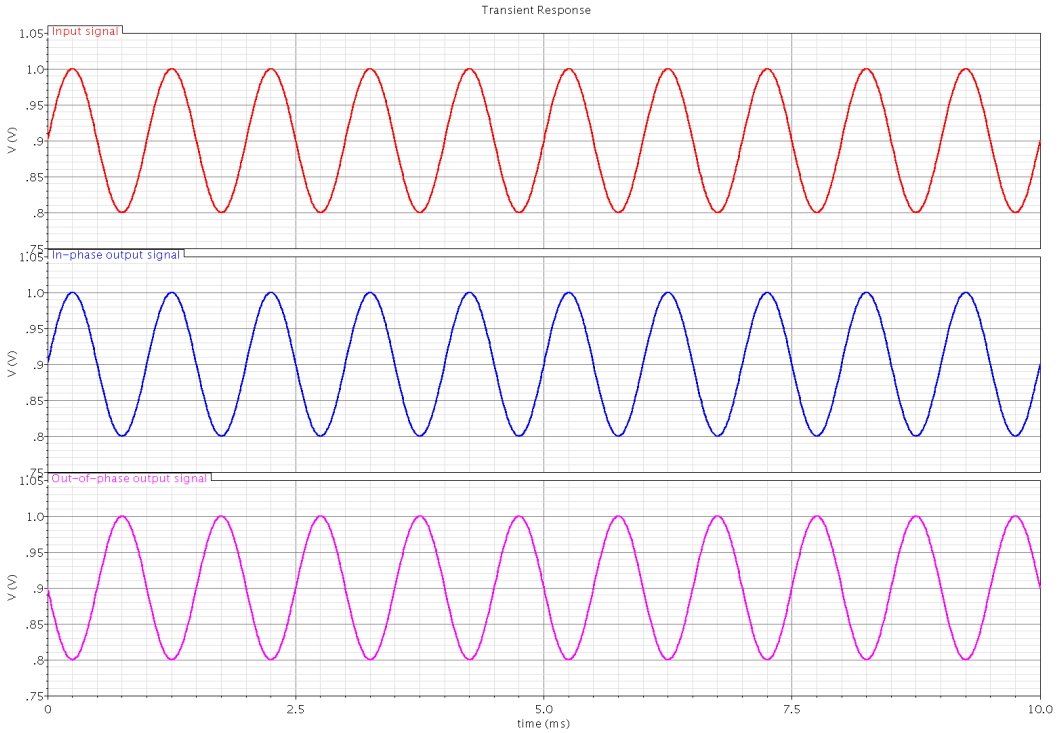

Figure 19: Plot showing the transient response of the single-ended to differential converter

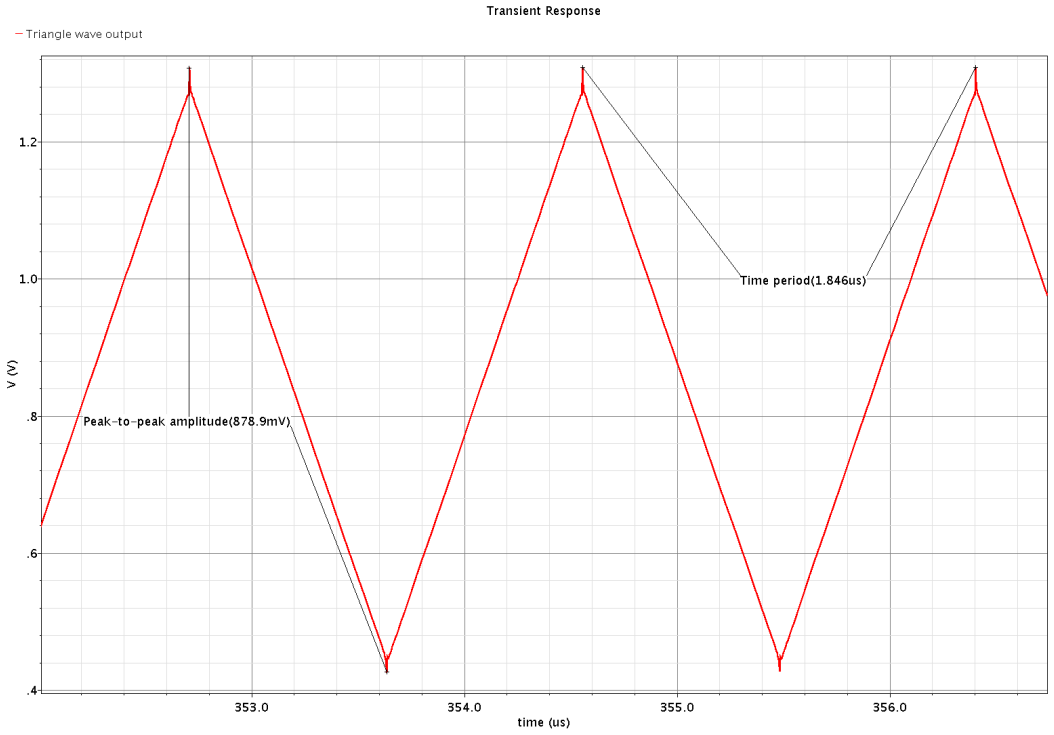

Figure 20: Plot showing the transient response of the triangle wave generator block

approximately 20ns to either edge of the input wave.

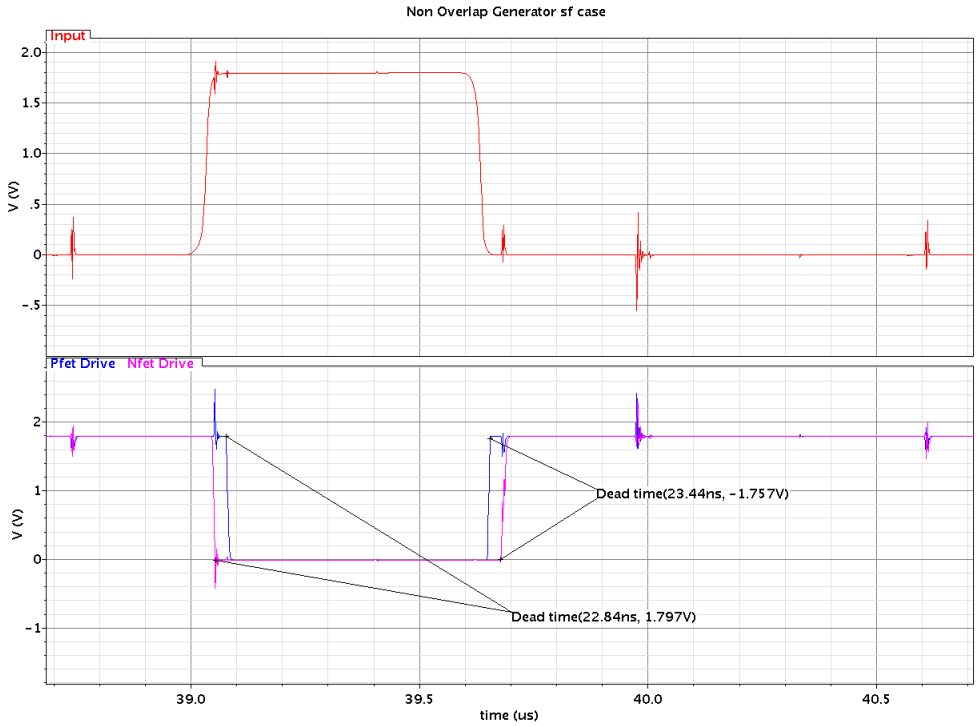

Figure 21: Plot showing the transient response of the non-overlap generator block (including buffers) in the SF case

4.6 System response

Figure 22 shows the individual and differential outputs of the entire system for an input sinusoid of 10kHz. The input wave changes sign at $50\mu s$ and the differential output follows suit. The current drawn through a supply with a 5Ω series resistance and $10\mu F$ decoupling capacitance is visible in the same plot.

Figure 23 shows the spectrum at the output and the input of the class-D amplifier. The SNDR in this case (input at approximately half the full-scale amplitude) is 33dB.

5 Test Strategy

This section discusses the various tests we have planned to test the correctness of functionality of different parts of the chip. Several points are tapped along the signal chain so that they can be observed. These points are connected through transmission gates to the pads. These transmission gates are controlled using an off-chip BuffEn signal so that the points being observed are not loaded by the pad parasitics during normal operation.

Characterization The pin used for supplying the bias current is connected to a diode-connected pFET. By varying the current being supplied to the pin and measuring the potential at the gate, we can obtain an I_D vs V_{SG} curve for the pFET. A test pin, BuffEn, which is used for enabling observation transmission gates on the chip is connected to a diode-connected nFET with the source of the transistor connected to V_{SS}. This pin can be used to similarly obtain the I_D vs V_{GS} curve for the nFET. Using these graphs, we can determine which process corner the chip falls under. Further, a $100k\Omega$ test resistor connected to the pin ExternalTriangle can be used to determine how much variation in resistance values can be expected on the chip.

Single-ended to differential converter The chip has separate V_{DD} pins for the analog and digital cir-

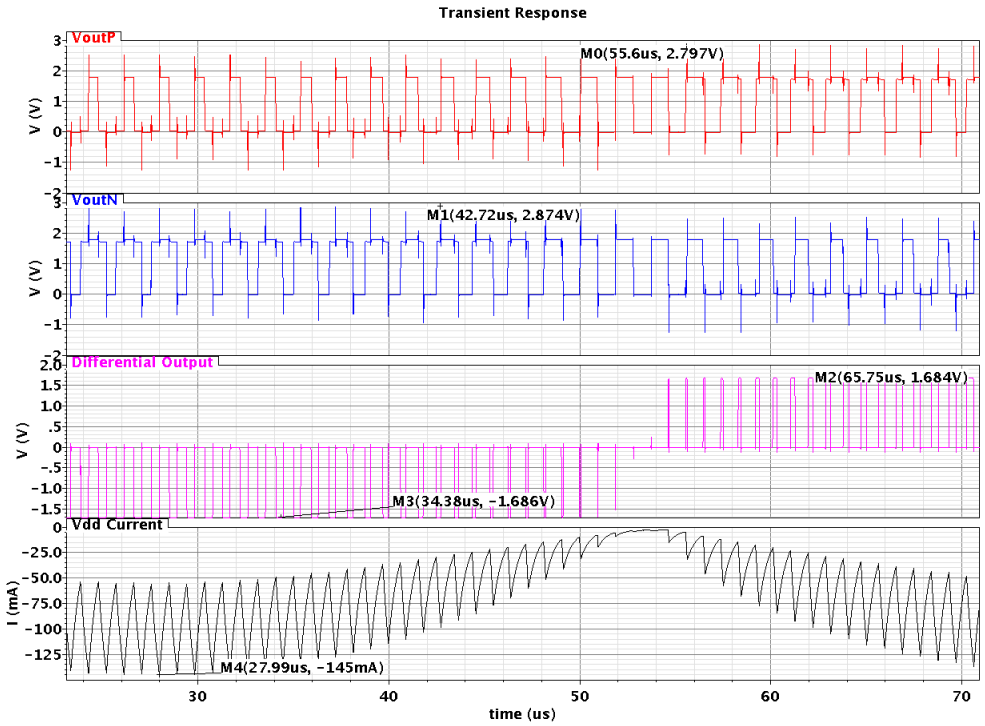

Figure 22: Plot showing the single-ended and differential output voltages and the current drawn from the supply

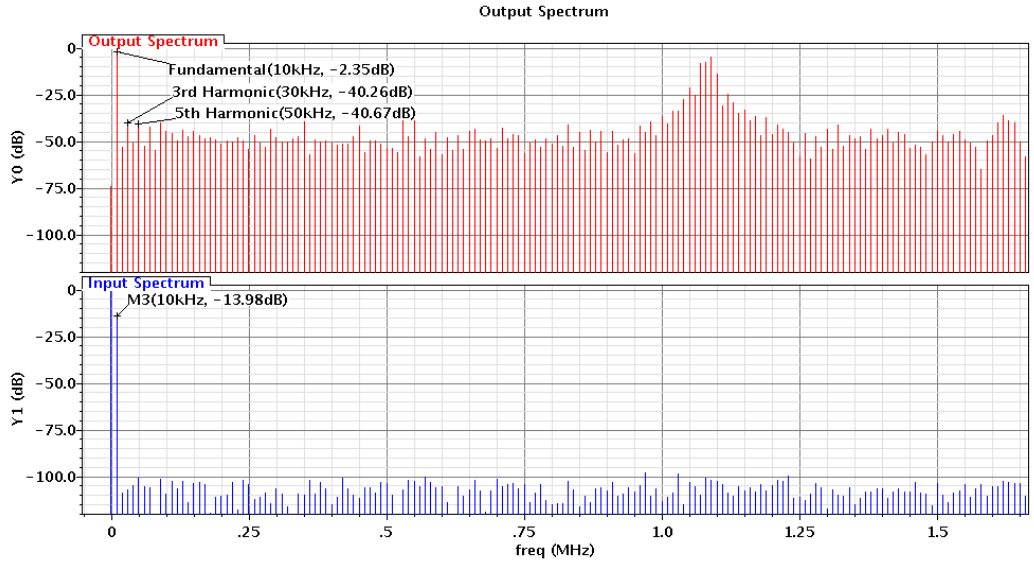

Figure 23: Spectrum at the input and output of the class-D amplifier for a 10kHz 200mV input

cuits. By powering off the digital circuits and observing the voltage at the pin through which we are supposed to close the loop, we can observe the output voltage of the single-ended to differential converter. In normal operation, the loop will be closed through an external resistor and the on-chip integrator will hold the potential of this pin close to V_{REF}.

Non-overlap Generator Both outputs of the non-overlap generator (after suitable buffering) are tapped so that the amount of non-overlap introduced by the circuit can be observed.

Triangle wave generator The pin ExternalTriangle controls whether the on-chip triangle wave generator or an externally provided triangle wave is connected to the PWM. If, for some reason, the

on-chip triangle wave generator does not behave as expected, this functionality enables us to operate the circuit using an external triangle wave which could be generated from a function generator. Further, we have also implemented the option to observe the output of the on-chip triangle wave generator.

6 Conclusion

This report presents some of the important choices made during the design of a first-order bridged three-level class-D amplifier. The circuit was initially designed to operate from dual supplies with the analog circuits operating at 1.8V and the digital circuits operating at 3.3V. Later, this was changed so that the entire circuit operated in the 1.8V power domain. Since the supplies are kept separate, it is still possible to independently turn on or off the digital and analog circuits.

Realizing the class-D amplifier included designing and laying out an OTA, a comparator, a non-overlap generator, several digital logic gates and the output driver transistors and finally putting all the individual blocks together. This was done using IBM's 180nm 7RF technology. The final layout including the circuitry and the bondpads was checked to be free of DRC and LVS errors. The amplifier is capable of providing 202.5mW to an 8Ω load.

A Pin Configuration

The pin configuration of the proposed amplifier is shown in Table 1. Each of the observation points are connected to the corresponding pads through a transmission gate. During debugging, the transmission gates are enabled so that the observation points in the circuit are connected to the pads. However, during normal operation, the transmission gates are disabled so that the circuit does not see the effects of the pad parasitics.

Section	S.No.	Pin Name	Pin Count	Comments
Supply Pins	1	PV_{DD}	3	Driven from high current limit 1.8V supply
	2	AV_{DD}	1	Driven from 1.8V supply
	3	V_{SS}	6	Analog and digital V_{SS} internally shorted
Observation Pins	4	Triangle wave	1	Check output of generated triangle wave
	5	Non overlap generator	2	Observe dead time and its effect on THD
	6	Comparator	1	Check PWM comparator output
Inout Pins	7	Output+	3	80mA per pin
	8	Output-	3	80mA per pin
	9	Input	1	Single ended input
	10	Bias current	1	$10\mu A$ current source
	11	Reference Voltage	1	Connect to nominal voltage of 0.9V
Bypass Pins	12	Triangle Bypass	2	1 for supplying triangle wave, one for controlling bypass
Control Pins	13	Buffer enable	1	Disable observation transmission gates during normal use
	14	Feedback	2	One pin per channel for closing the feedback loop
		Total pins	28	

Table 1: Pin Configuration

B Top-level schematic and pin layout

Figure 25 shows the top-level schematic as made using Virtuoso while Figure 24 shows the PCB level connections. Note that this view does not include the bondpads, which are present in the view above the one shown. Figure 26 shows the pin-out of the chip.

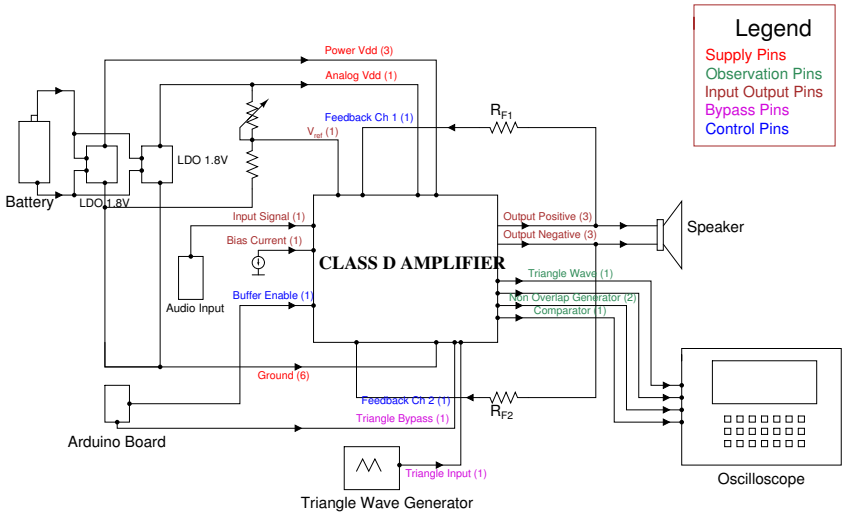

Figure 24: Typical PCB level connections

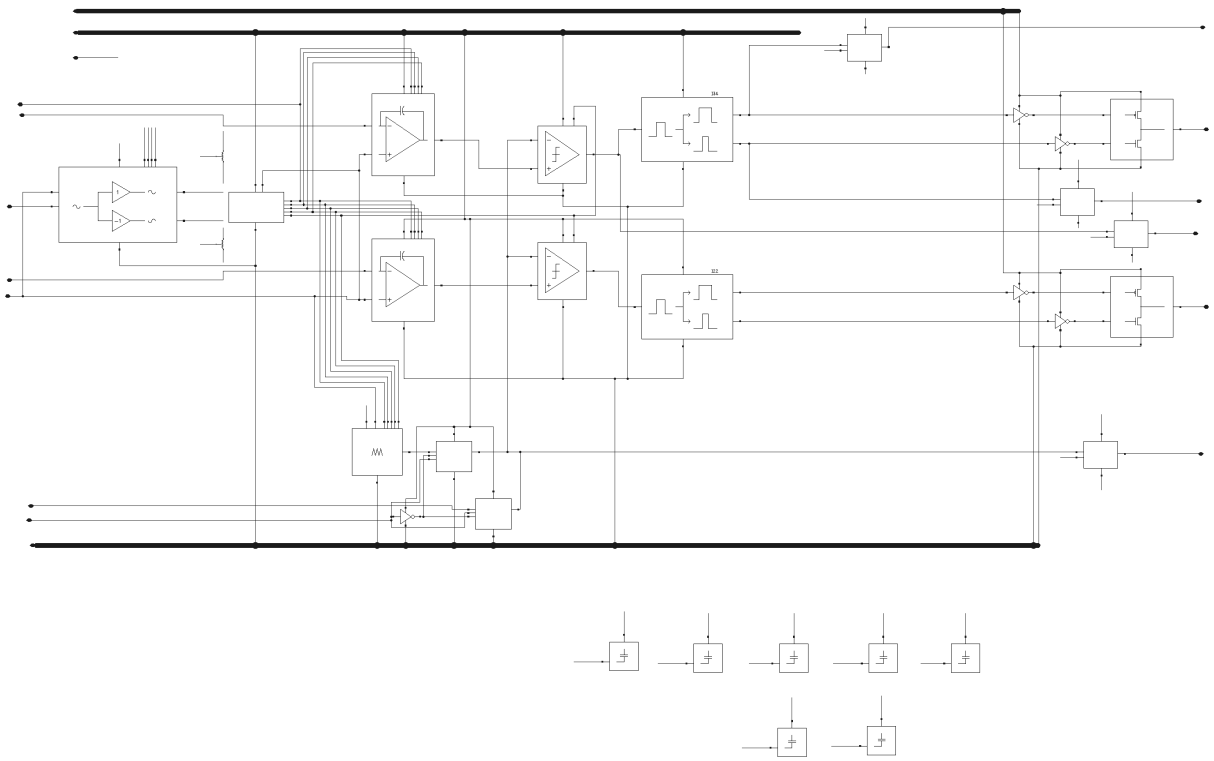

Figure 25: Top-level schematic view of the class-D amplifier as designed using Virtuoso

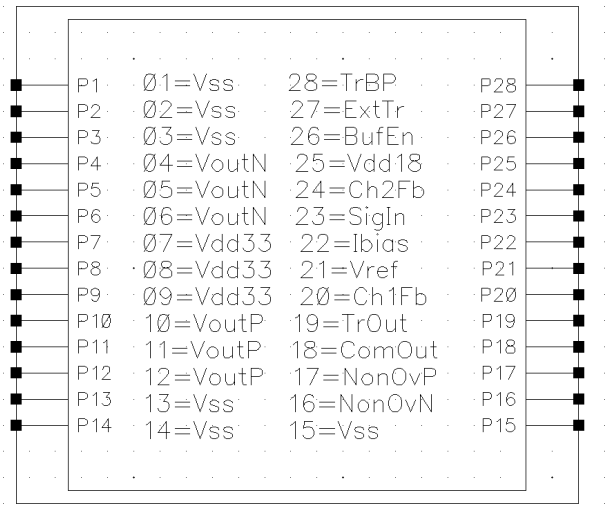

Figure 26: Pin-out of the chip