BRNO UNIVERSITY OF TECHNOLOGY

VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

FAKULTA ELEKTROTECHNIKY
A KOMUNIKAČNÍCH TECHNOLOGIÍ

DEPARTMENT OF MICROELECTRONICS

ÚSTAV MIKROELEKTRONIKY

DESIGN OF LOW ORDER HIGH OSR DISCRETE TIME DELTA-SIGMA MODULATOR FOR AUDIO APPLICATIONS

NÁVRH DISKRÉTNÍHO DELTA-SIGMA MODULÁTORU PRO AUDIO APLIKACE NÍZKÉHO ŘÁDU S VYSOKÝM KOEFICIENTEM PŘEVZORKOVÁNÍ

MASTER'S THESIS
DIPLOMOVÁ PRÁCE

AUTHOR — **Bc. Jaroslav Dohnal**
AUTOR PRÁCE

SUPERVISOR — **doc. Ing. Jiří Háze, Ph.D.**
VEDOUCÍ PRÁCE

BRNO 2020

BRNO UNIVERSITY OF TECHNOLOGY
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

Master's Thesis

Master's study field **Microelectronics**

Department of Microelectronics

Student: Bc. Jaroslav Dohnal ***ID:*** 174198

Year of study: 2 ***Academic year:*** 2019/20

TITLE OF THESIS:

Design of low order high OSR discrete time delta-sigma modulator for audio applications

INSTRUCTION:

Design and simulate a low order delta-sigma modulator of a chosen architecture in a discrete time domain that will operate in an audio band with ENOB of at least 18 bits. Implement the modulator in a FPGA together with the interpolation filter and verify its properties on the prototype.

RECOMMENDED LITERATURE:

According to recommendations of supervisor

Date of project specification: 3.2.2020 ***Deadline for submission:*** 1.6.2020

Supervisor: doc. Ing. Jiří Háze, Ph.D.

doc. Ing. Lukáš Fujcik, Ph.D.
Subject Council chairman

ABSTRACT

This master thesis aims to familiarize the reader with the basic concept and fundamental principles of single-loop ΔΣ modulators. It offers an alternative to a high-order ΔΣ modulators in the form of low-order ΔΣ modulators running at high oversampling rate. Low order modulators have better modulator loop stability, which can be leveraged to get higher noise-shaping power at lower frequencies. A complete ΔΣ digital to analog converter is proposed, mostly implemented in an FPGA. A hardware prototype was built to evaluate the ΔΣ DAC implementation.

KEYWORDS

delta-sigma modulator, converter, DAC, FPGA, audio, oversampling

ROZŠÍŘENÝ ABSTRAKT

Diplomová práce ozřejmuje čtenáři problematiku delta-sigma (ΔΣ) modulátorů s jednou zpětnovazební smyčkou. Zabývá se základními principy převzorkování u číslicově-analogových převodníků a rozšiřuje je o teorii tvarování spektra šumu. Popsány jsou také kvantizéry a zdůrazněny nevýhody jednobitového kvantizéru, zejména nemožnost jeho linearizace bílým šumem [12].

Vycházeje z této teorie jsou navrženy tři jednosmyčkové ΔΣ modulátory v podobě druhého řádu, třetího řádu a pátého řádu, které běží na 1024 OSR jako alternativa k běžně používáným modulátorům vysokých řádů. Tyto modulátory nižších řádů přináší benefity v podobě lepší stability smyčky bez nutnosti použití mnoha komplexních pólů filtru smyčky tvarování šumu. Modulátory jsou podrobeny testům v prostředí MATLAB Simulink, kde dosahují ENOB 21,9 bitů pro druhý řád, 27,7 bitů pro třetí řád a 37,1 bitů pro pátý řád modulátoru v audio pásmu. Jádra těchto modulátorů jsou přímo zkompilována do VHDL pomocí MATLAB HDL coderu. K jádrům jsou ve VHDL dopsány vstupně-výstupní porty, generátory náhodných čísel a 2,2-bitový (pěti-úrovňový) kvantizér, uzavírající smyčku modultátoru.

ΔΣ modulátory jsou doplněny interpolačním filtrem kladoucí důraz na korektní rekonstrukci časové informace v audio signálu. Toho je dosaženo interpolačním filtrem s nezvykle vysokým počtem 32000 koeficientů. Tyto koeficienty jsou tvořeny diskretizovanou, téměř ideální sinc funkcí aby došlo k co nejlepší rekonstrukci signálu dle Nyquist-Shannonova theorému o rekonstrukci diskretizovaného signálu. Filtr je navrhnutý okenní metodou nad ideální sinc funkcí. Byla navržena speciální okenní funkce jako kompromis mezi časovou a frekvenční odezvou rekonstrukčního filtru. Filtr interpoluje vstupní signal 16× a je tak schopný rekonstruovat hudební signal až do výsledného časového rozlišení, které je lepší než 4 µs v úseku 20 ms signálu (skupinové zpoždění interpolátoru). Toto časové rozlišení 4 µs je limit, kdy lidský sluch je ještě schopný rozlišit rozdíl ve zpoždění signálu mezi levým a pravým uchem [9] [10]. Mozek dle tohoto rozdílu pak tvoří zásadní obraz o svém okolí, tj. hudební scéna je daleko lépe zrekonstruována.

Rekonsturkční filtr je doplňen vysokofrevenčním filtrem, jež odstraňuje spektrální obrazy signálu až do cca 6 MHz. Společně jsou tyto modulatory a filtry implementovány do FPGA Artix 7 35T výrobce Xilinx.

K otestování vlastností vytvořeného ΔΣ číslicově-analogového převodíku byly navrženy desky plošných spojů pro diskrétní převodník 2,2bitového, pulsně-hustotně kódovaného signálu z modulátorů na analogový signál. Tento převodník byl realizován jako čtyřprvkový pětiúrovňový zdroj proudu, schopný provozu na 100 MHz. Každý prvek je tvořen dvojicí klopných obvodů typu D a čtyř precizních rezistorů na každém z výstupů těchto klopných obvodů. Lienarizace číslicově-analogového převodníku byla docílena náhodným přepínáním prvků mezi sebou, čímž je docíleno přeměny absolutní nepřesnosti (nesouladu, angl. mismatch) každého elementu do podoby mírně zvýšeného šumového dna.

Diskrétní operační zesilovač se zvýšenou odolností na elektromagnetické vysokofrekvenční rušení pro výstupní dolnopropustný filtr je v rámci práce taktéž navržen. Byla zvolena architaktura složené kaskody s JFET vstupním diferenciálním zesilovačem. JFET tranzistory byly zvoleny z důvodu jejich zvýšené iminuty na vysokofrekvenční rušení (nedochází k tak silné demodulaci jako na PN přechodu bipolárních tranzistorů). Dva vstupní diferenciální páry navíc pracují paralelně, čímž je docíleno nižšího šumu a lepšího souběhu. OZ byl realizován zejména pomocí dvojitých diskrétních tranzistorů zlepšující souběh a zabírající méně místa na DPS.

Navržený operační zesilovač tvoří jádro dolnopropustného výstupní filtru integrující hustotně modulované proudové pulzy. Byla zvolena Bessel aproximace filtru z důvodu konstatního skupinového zpoždění v propustném pásmu. Výstupní filtr pracuje jako diferenciální sumační zesilovač. Tím pádem pozitivní fáze pulzního převodníku nepracuje do nulové impedance (v proudovém modu), způsobující nerovnoměrnost mezi pozitivní a negativvní větví DACu. Tyto nerovnoměrnosti se projevují jako vysoká druhá harmonická složka při měření harmonického zkreslení. Lepším řešením by byl plně diferenciální operační zesilovač.

Celkově bylo však dosaženo 0.003 % harmonického zkreslení, téměř identického pro všechny tři modulátory. Takto vysoké zkreslení je způsobeno diferenciálním nesouladem zmíněným výše. Taktéž SNR a DR bylo dosaženo 119 dBr pro všechny tři modulátory.

Tato skutečnost dokazuje, že i s modulátorem nízkého druhého řádu lze dosáhnout dostatečných parametrů i pro náročnou reprodukci audio signálu. Nutnou podmínkou je však vysoký faktor převzorkování. Přehledné shrnutí výsledků je také v Tabulce 16 v závěrečné kapitole.

KLÍČOVÁ SLOVA

delta-sigma modulátor, číslicově-analogový převodník, DAC, FPGA, audio, převzorkování, pulsně-hustotní modulace

DOHNAL, Jaroslav. *Design of low order high OSR discrete time delta-sigma modulator for audio applications.* Brno, 2020. Also available from: https://www.vutbr.cz/studenti/zav-prace/detail/126936. Master thesis. Brno University of Technology, Faculty of Electrical Engineering and Communication. Department of Microelectronics.

ACKNOWLEDGMENT

I would like to thank my supervisor, doc. Ing. Jiří Háze, Ph.D., for the patient guidance and advice he has provided throughout my time as his student at the Department of Microelectronics. I also want to thank John A. Westlake for the help with the analog PDM DAC section and the encouragement he provided.

TABLE OF CONTENTS

INTRODUCTION

Today, most audio *digital to analog converters*, or *DAC*, rely solely on *ΔΣ (delta-sigma) modulators*. ΔΣ modulation based DACs offer good performance and low power consumption. ΔΣ DACs can also be implemented on a chip more easily compared to their multi-bit predecessors. However, most of them are implemented using high order modulators with complex poles to ensure performance and stability. These may sound worse as they may be prone to noise floor modulation and other undesirable effects. Listening tests have also shown that the resolving power is needed most at lower frequencies [5].

Alternative low order modulators are presented in this thesis. They achieve similar performance as high-order modulators by leveraging a high oversampling rate. No resonator sections need to be used to achieve stability, leaving most of the decorrelating power at the lower end of the audio band. The ΔΣ modulators are implemented in an FPGA. The modulators *pulse-density modulated* output pulses are converted into an analog signal by an *analog low pass filter*.

Interpolating reconstruction filter is designed with an emphasis on time-domain reconstruction performance. Interpolation filters and other support blocks are also implemented in the FPGA.

This thesis presents a complete ΔΣ DAC solution, including an interpolation filter, a ΔΣ modulator, and an analog section. Sample rates of 44.1 kHz and 48 kHz are supported at this time. However, an extension to 768 kHz is possible by modifying the reconstruction filter state machine.

1 DELTA-SIGMA CONVERTERS

Delta-Sigma (ΔΣ) converters can achieve very high resolution at low frequency by utilizing the *oversampling* process combined with the *noise shapin*g technique. Therefore they are well suited for audio applications as the dynamic range of human hearing is limited at high frequencies (above 15 kHz). This chapter is a brief introduction into basic general concepts of ΔΣ converters.

1.1 Oversampling and Noise Shaping

An *ADC* example is used to demonstrate oversampling and noise shaping concepts. The concepts are identical for ΔΣ DACs, however, the input signal is a discrete-time signal rather than a continuous-time signal.

The signal has to be quantized by a *quantizer* to distinct, equally spaced, voltage levels in order to be processed in the discrete-time domain. Quantizer is a non-linear system and introduces *quantization error* when approximating the input signal. Analyzing this error can be quite complex [1], hence it is usually approximated as white noise. This noise is then referred to as *quantization noise*.

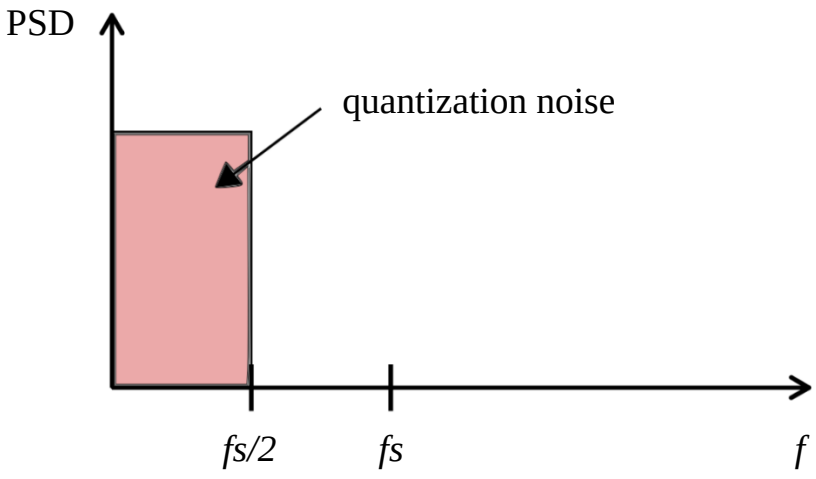


Figure 1 - Spectral density of quantization noise in a Nyquist system

Figure 1 shows the essential Nyquist system sampling at the *sampling rate fs*. The quantization noise is distributed uniformly within the *Nyquist band* from DC to

$$BW_{qnoise} = \frac{f_s}{2}, \tag{1.1}$$

where f_s is the sampling rate [1]. Spectral power density is defined as

$$P_e(f) = \frac{V_q}{\sqrt{12 f_s}}, \tag{1.2}$$

where V_q is the quantizer error [1], [2]. *Signal to Noise Ratio* (*SNR*) is then defined as

$$SNR_{max} = 6.02\ ENOB + 1.76, \tag{1.3}$$

for full-scale harmonic input, where *ENOB* is an *Effective Number Of Bits* of the converter [1].

This systems SNR can be increased in three ways

- increasing the number of bits (levels) of the quantizer,
- increasing sampling speed at a constant bandwidth – oversampling,
- shifting the noise out of the band of interest – noise shaping.

The case for oversampling is shown in Figure 2. The noise is uniformly distributed at the bandwidth of

$$W_{qnoise} = OSR \times f_s 2, \tag{1.4}$$

where *OSR* is the *Over Sampling Ratio*. Noise is removed out of the band when a *Nyquist bandwidth* filter is applied.

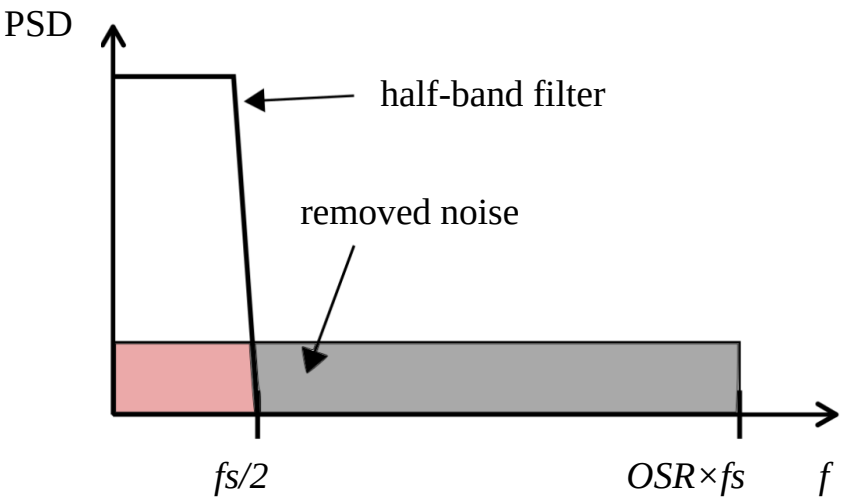


Figure 2 - Spectral density of quantization noise in an oversampling converter

The SNR of an oversampled converter corresponds to

$$SNR_{max} = 6.02\ ENOB + 1.76 + 10\log(OSR). \quad (1.5)$$

The first part of this equation is identical to the equation (1.3). The second part expresses the increased SNR due to OSR. The OSR has to be equal to 4 to increase ENOB by one bit. In other words, oversampling results in a 0.5-bit increase per each oversampling ratio doubling.

It is impractical to increase resolution only by oversampling. High-resolution systems can be realized by combining oversampling with noise-shaping while keeping a reasonable OSR ratio. Noise-shaping moves most of the quantization noise out of the band of interest. For example, an audio noise-shaping converter moves most of the noise from the audio band into a high-frequency noise where it is removed subsequently by an audio bandwidth filter.

This situation is shown in Figure 3. The key concept behind all ΔΣ converters is the use of oversampling and noise-shaping.

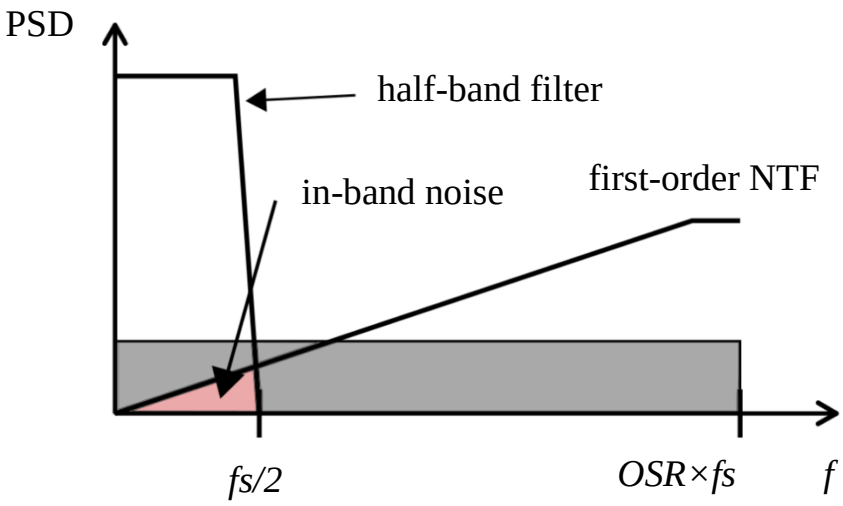


Figure 3 - spectral density of quantization noise in a noise shaping ΔΣ converter

1.2 Delta-Sigma Modulators

Delta-Sigma modulators consist of three basic blocks [1]

- loop filter,
- quantizer,
- feedback DAC or DDC.

The block diagram in Figure 4 shows a ΔΣ modulator structure. The *U* is the input signal, *Y* is the loop filter output, *E* is the quantization error, i.e., quantization noise, and *V* is the modulator output. The DAC (for continuous-time) or *DDC* (*Digital to Digital Converter* – for the discrete-time domain) scales the quantizer output to the same format as the input signal *U*.

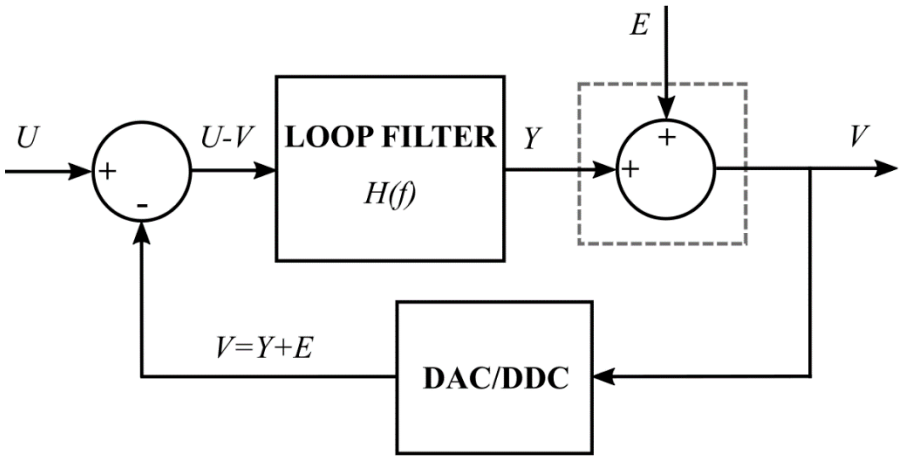


Figure 4 - A typical ΔΣ converter block diagram

Frequency domain analysis of the Figure 4 loop filter yields a loop transfer function [1] of

$$H(f) = \frac{1}{f} \tag{1.6}$$

which is an integrator transfer function, typical for a first-order loop filter. The amplitude response *H(f)* is inversely proportional to frequency *f*, resembling a low-pass filter when the loop is closed. The expression for the output signal *V* can be written as

$$V = \frac{1}{f}\,(U - V) + E. \tag{1.7}$$

Solving for *V* is then

$$V = \frac{U}{f+1} + \frac{Ef}{f+1}. \tag{1.8}$$

The first term is a *Signal Transfer Function (STF)*, and the second term is a *Noise Transfer Function (NTF).* According to these transfer functions, as the frequency *f* approaches zero, the STF approaches *U* as the NTF approaches zero. Hence the output signal *V* is the input signal *U* with no noise component. The output *V* mainly consists of quantization noise *E* at high frequencies. Therefore, the desired noise shaping characteristics are determined by the loop filter.

A loop filter is usually built with one or more cascaded integrators that are connected with feedback and feedforward paths in high-order systems. The number of integrators determines the ΔΣ modulator order. The modulator order is directly associated with the noise shaping characteristics as well as the modulator stability. Single loop ΔΣ modulators with more than two integrators are inherently unstable and must be compensated [1].

The noise shaping characteristics of first, second, and third-order ΔΣ modulators are illustrated in Figure 5. The third-order NTF is superior in moving the quantization noise out of the band. However, it is more difficult to design due to its higher order. ASICs with no more than sixth-order modulators are usually used in practice [2], but cases with 17th order modulators can be found [3].

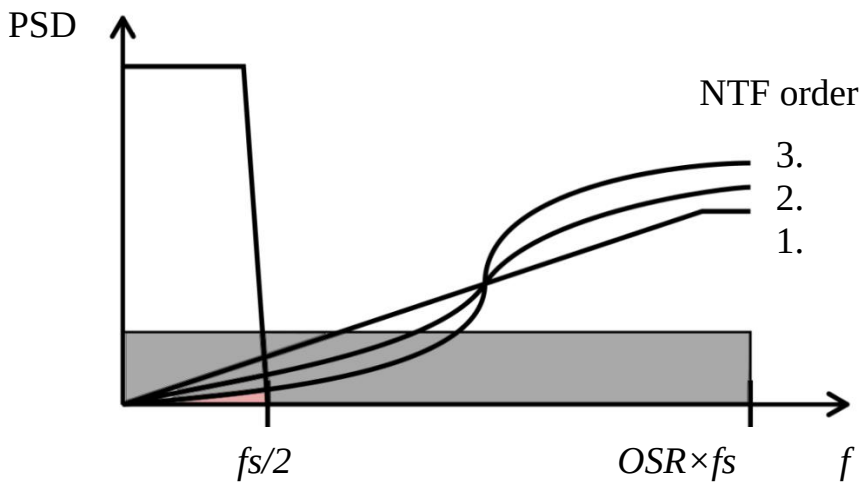


Figure 5 - Noise shaping characteristics of different order ΔΣ modulators

ΔΣ converter maximum SNR corresponds to

$$SNR_{max} = 6.02\,ENOB + 1.76 + (20L + 10)\log(OSR) - 10log\frac{\pi^{2L}}{2L+1}, \quad (1.9)$$

where L is the order of the modulator [2]. The SNR is lower in real implementations, as compromises between the modulators performance and stability have to be made.

A ΔΣ modulator quantizer output is a *Pulse Density Modulated* (PDM) signal. This signal is represented by a ratio of zeros and ones in a specified time period. PDM signal can be represented with one or more bits, but no more than 6-bit quantizers are usually used. How well the noise-shaping modulators can correct for a 1-bit quantizer quantization error is shown in Table 1.

Table 1 - Theoretical achievable SNR for 1-bit converters

	SNR [dB]			
OSR	OSR Conv.	1st order ΔΣ	2nd order ΔΣ	3rd order ΔΣ
256	31.9 dB	74.9 dB	115.3 dB	155 dB
512	34.9 dB	83.9 dB	130.3 dB	176 dB
1024	37.9 dB	92.9 dB	145.4 dB	197 dB

Table 1 compares 1-bit converters running at the three high oversampling ratios, which are still reasonably achievable. Oversampling converters achieve SNR increment of 3 dB/octave. A first-order noise-shaping achieves 9 dB/octave. The second-order adds 6 dB resulting in 15 dB/octave while the third-order achieves 21 dB/octave. Each quantizer bit will also add about 6.02 dB to absolute SNR values.

1.3 Quantization

As mentioned in Chapter 1.1, quantization is a process of converting the amplitude of continuous or higher resolution discrete signal to a lower resolution discrete signal. Quantization is usually uniform, which means that any two adjacent quantized values differ by a fixed level spacing Δ. Quantizer has an input *Y* and an output *V* and is defined by its *Y-V transfer curve* [4].

Figure 6 shows a quantizer transfer curve example for a symmetric bipolar quantizer with four levels (steps). In this case, $Y = 0$ coincides with a step rise of *V* and hence is called a *mid-rise quantizer*. It is often desirable to approximate the quantizer transfer curve with a line $V = Yk$, where k is the gain of the quantizer. Deviation from this

ideal characteristic is called the quantization error or, not entirely correctly [4], the quantization noise. The quantization noise is shown in Figure 6B. Coefficient $k = 1$ was used for the approximation in all examples.

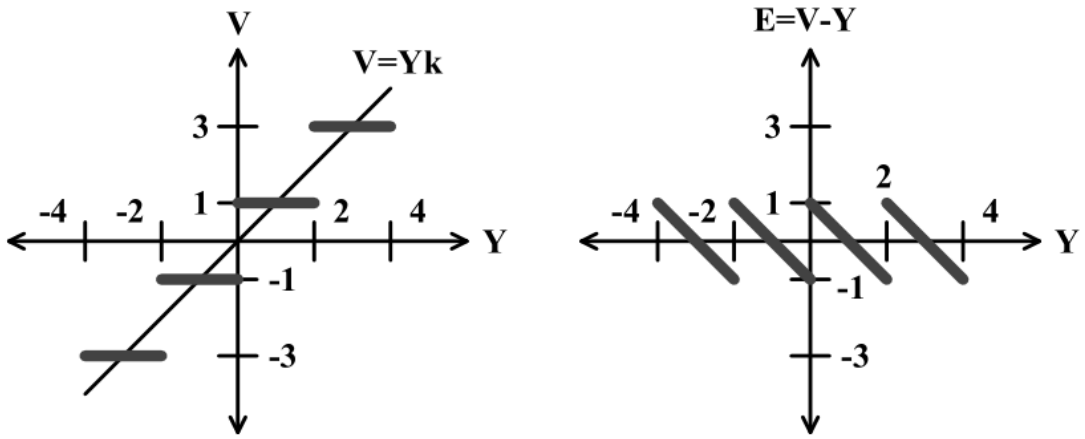


Figure 6 - Transfer and error curves of a symmetric 4-step mid-rise quantizer

Figure 7 shows a similar multilevel symmetric bipolar quantizer, but with five steps. Point $Y = 0$ now occurs in the middle of the flat portion (tread), and hence the quantizer is called a *mid-tread* quantizer. The variable determining the number of quantizer steps is usually called *M*. Quantizers are mid-rise for even-*M* of steps and mid-thread for odd-*M* steps.

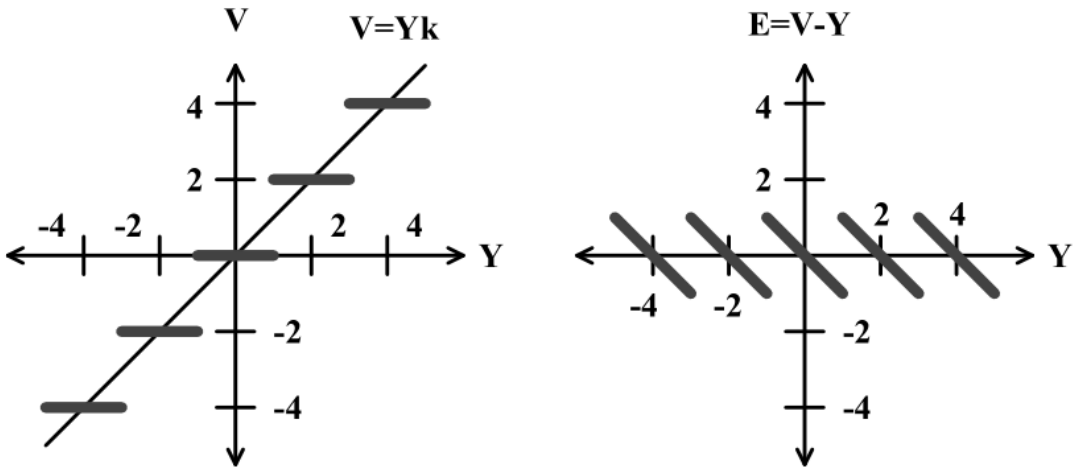


Figure 7 - Transfer curve and error of a symmetric 5-step mid-tread quantizer.

The simplest quantizer possible is shown in Figure 8. This quantizer only has two steps. The maximum output value occurs when $Y > 0$ and the minimum output value occurs when $Y \leq 0$. A typical implementation of this quantizer is a 1-bit comparator with zero. This quantizer is used in 1-bit ΔΣ converters. A SNR of 176 dB is achievable with this quantizer and the third-order ΔΣ modulator running at OSR of 512 (Table 1).

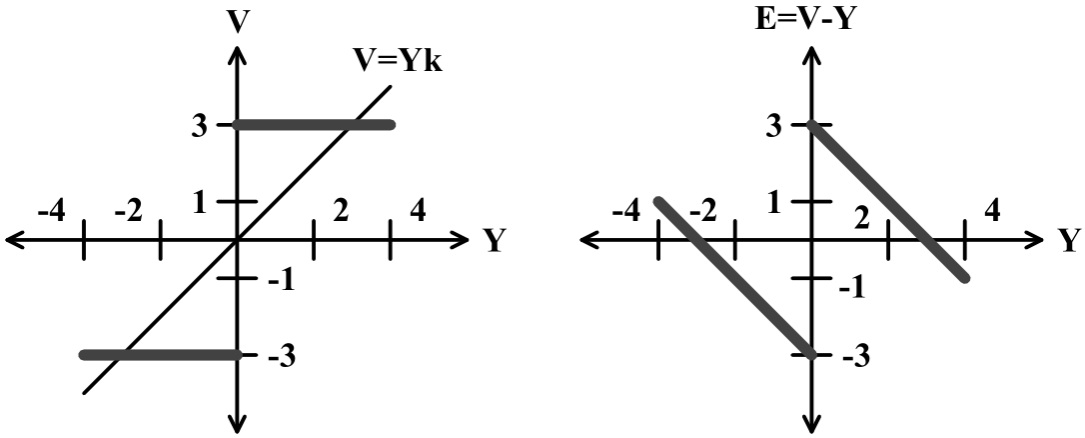


Figure 8 - Transfer curve and error of a 2-step quantizer.

Simplicity and linearity are the main advantages of 1-bit quantizers. Nevertheless, ΔΣ modulators are more prone to instabilities and idle tones when designed around 1-bit quantizers. This thesis uses a *multi-bit* quantizer from Figure 7. It offers a compromise between modulator stability, dynamic range, and linearity. The method to increase the linearity of these multi-bit DACs is discussed in Chapter 3.

All presented quantizers are symmetric bipolar quantizers as audio signals use signed number representation. *Unipolar* quantizers also exist, but they are not used very often in audio applications.

1.4 Idle tones

The ratio between zeros and ones rises (long series of ‘1’ and ‘0’) at the ΔΣ modulator output when the input signal of the modulator gets close to the feedback value (i.e. there is little to no error). This causes a cyclic behaviour of the modulator, creating higher harmonic and non-harmonic tones called *idle tones*, which are often audible. The quantization noise also gets correlated to the input signal. [2]

Idle tones can be suppressed by a white noise injection between the last integrator and the quantizer. This noise forces the ΔΣ modulator to correct the error within the loop continually, and the constant level on the output [2] is reduced. This process is sometimes called *modulator dithering*. Usually, a dither level around 0.5 LSB is used, but a higher level of dither might be needed for low order modulators. Paper [12] shows that the quantizer can be linearized entirely by an injection of ±1 LSB TPDF dither. Idle tones disappear entirely with a linear quantizer. However, the feedback has to correct for the added dither, and hence the modulator SNR is lowered. Modulators working in the continuous-time domain can be dithered with thermal noise [2]. Dither can be created with pseudo-random generators for discrete-time domain modulators.

1.5 ΔΣ Modulator Architectures

There are many different discrete-time ΔΣ modulator architectures. Only the simpler single loop architectures are considered in this thesis.

Some modulator architectures achieve higher SNR by adding or manipulating zeros and poles of the NTF and placing them in the passband using resonator sections. The result is increased SNR at higher frequencies at the cost of low-frequency noise shaping power. The Implemented modulators have these resonator sections multiplexed so that they can be disabled to evaluate the benefit of complex poles.

1.5.1 Single loop first-order modulator

The simplest ΔΣ modulator is shown in Figure 9. It is an example of a single loop first-order ΔΣ modulator with one non-delaying integrator. The integrator is created by a unit delay element (usually D-type flip flop clocked at sample rate frequency) and an adder. The output quantizer (multilevel or single level comparator) determines the output bit depth of the modulator. The feedback loop delay is needed to ensure stability and represents time to quantize, sample and feedback the integrator.

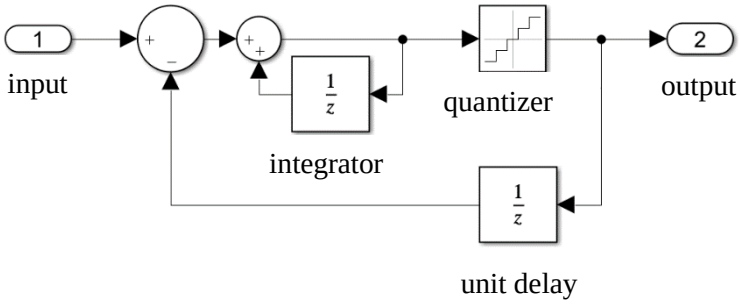


Figure 9 – A first-order ΔΣ modulator example

This first-order modulator is the simplest ΔΣ modulator possible and is very rarely used as a single modulator in a converter. It may be used where high step count quantizers are employed or in simple applications such as a DAW element matching block for mismatch error noise-shaping [1]. It is very susceptible to generating idle tones and has low noise-shaping power. On the other hand, the loop of this modulator is usually stable.

1.5.2 Single loop second-order modulator

Another integrator is added to the loop of the first-order modulator to create a second-order ΔΣ modulator, as shown in Figure 10.

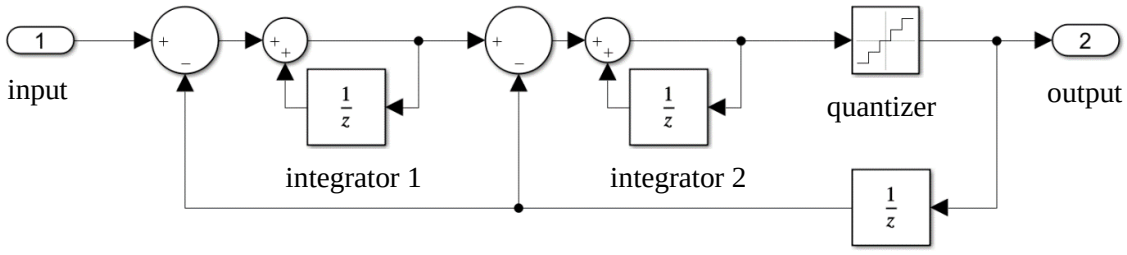


Figure 10 – A second-order ΔΣ modulator example

The second-order modulator is more resilient to cyclic behaviour as two cascaded integrators act more chaotically and thus the modulator does not settle too quickly [6]. Dither addition is needed to achieve sufficient performance for audio applications.

Figure 11 shows an alternative second-order architecture, sometimes called *Bosey-Wooley*. This ΔΣ modulator uses delaying integrators, hence no feedback delay is needed to ensure stability. The delaying integrators also help to pipeline the modulator [1]. Stability is further improved with gain blocks *A* and *C*. The gain block coefficients can shift poles out of band, while zeros remain at DC.

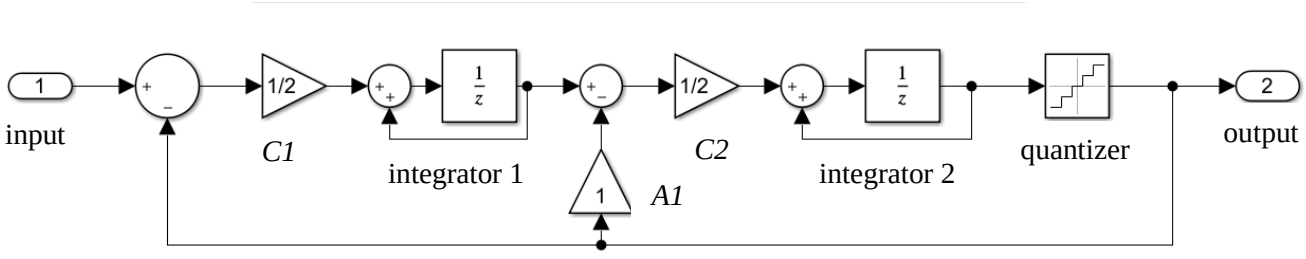


Figure 11 - Bosey-Wooley ΔΣ modulator

Similarly, more integrator blocks and feedback/feedforward paths can be added to create higher modulator orders. The next chapter briefly discusses the most popular generalized ΔΣ modulator architectures.

1.5.3 Generalized single-loop architectures

There are four main ΔΣ modulator loop architectures

- cascade-of-integrators, feedback form (CIFB),
- cascade-of-resonators, feedback form (CRFB),
- cascade-of-integrators, feedforward form (CIFF),
- cascade-of-resonators, feedforward form (CRFF).

CIFB ΔΣ modulator architecture example is shown in Figure 11.

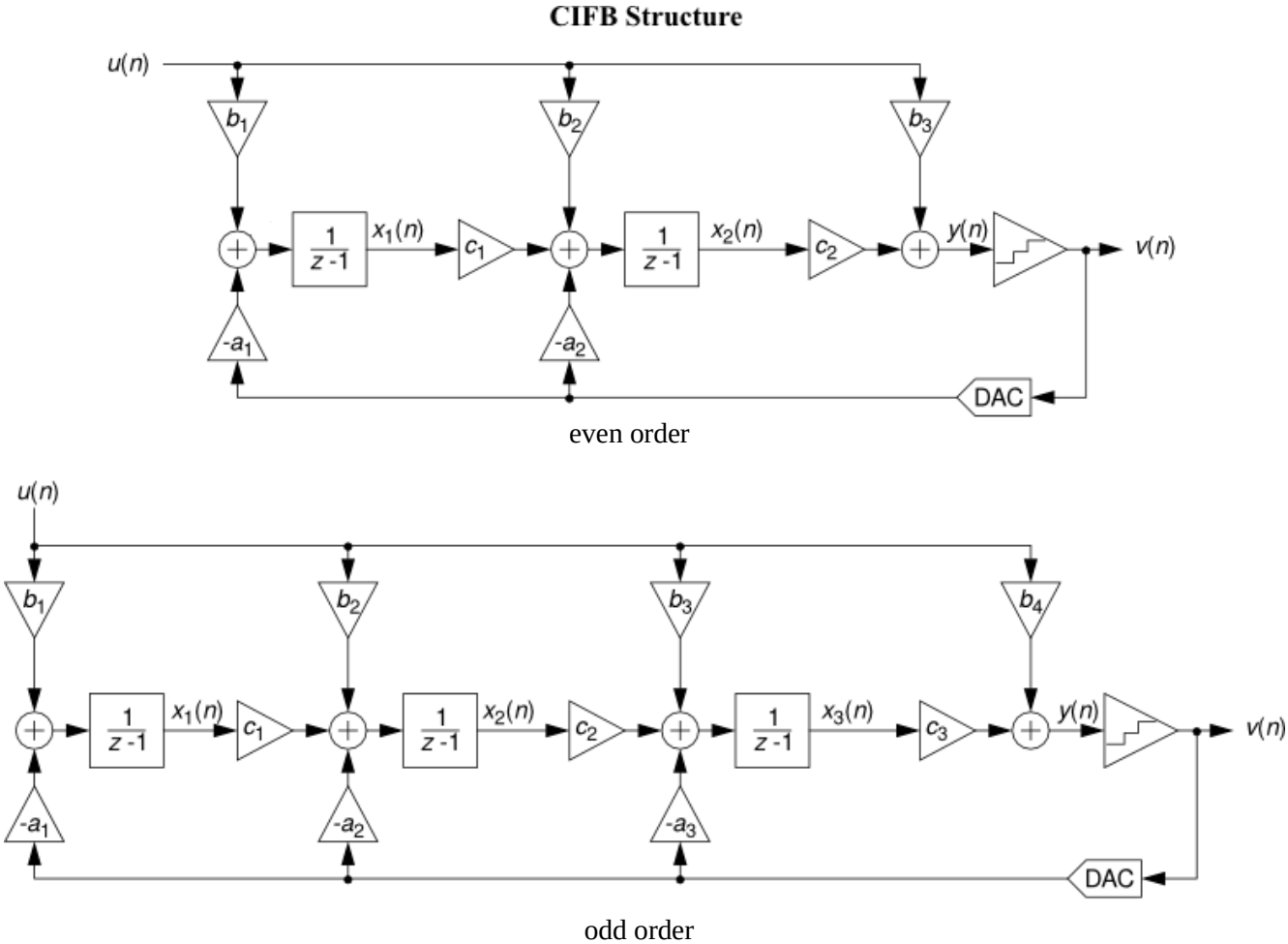


Figure 12 - CIFB ΔΣ modulator architecture example [14]

Signal feed gain blocks *B* are added to the modulator from Figure 11. Each of the *A*, *B* or *C* coefficients have a gain value between 0 and 1. The block is omitted entirely when the coefficient is 0. Similarly, the block is replaced with a short for the gain value of 1. Coefficients are ideally designed as a multiple of $1/2^N$. Simple bit-shifts can then be used to implement the division operation instead of implementing slower and more complex hardware multipliers.

Resonator sections are added to create the CRFB architecture. These resonator sections can extend the modulators SNR by creating zeros in the NTF passband. Generalized CRFB architecture is shown in Figure 13.

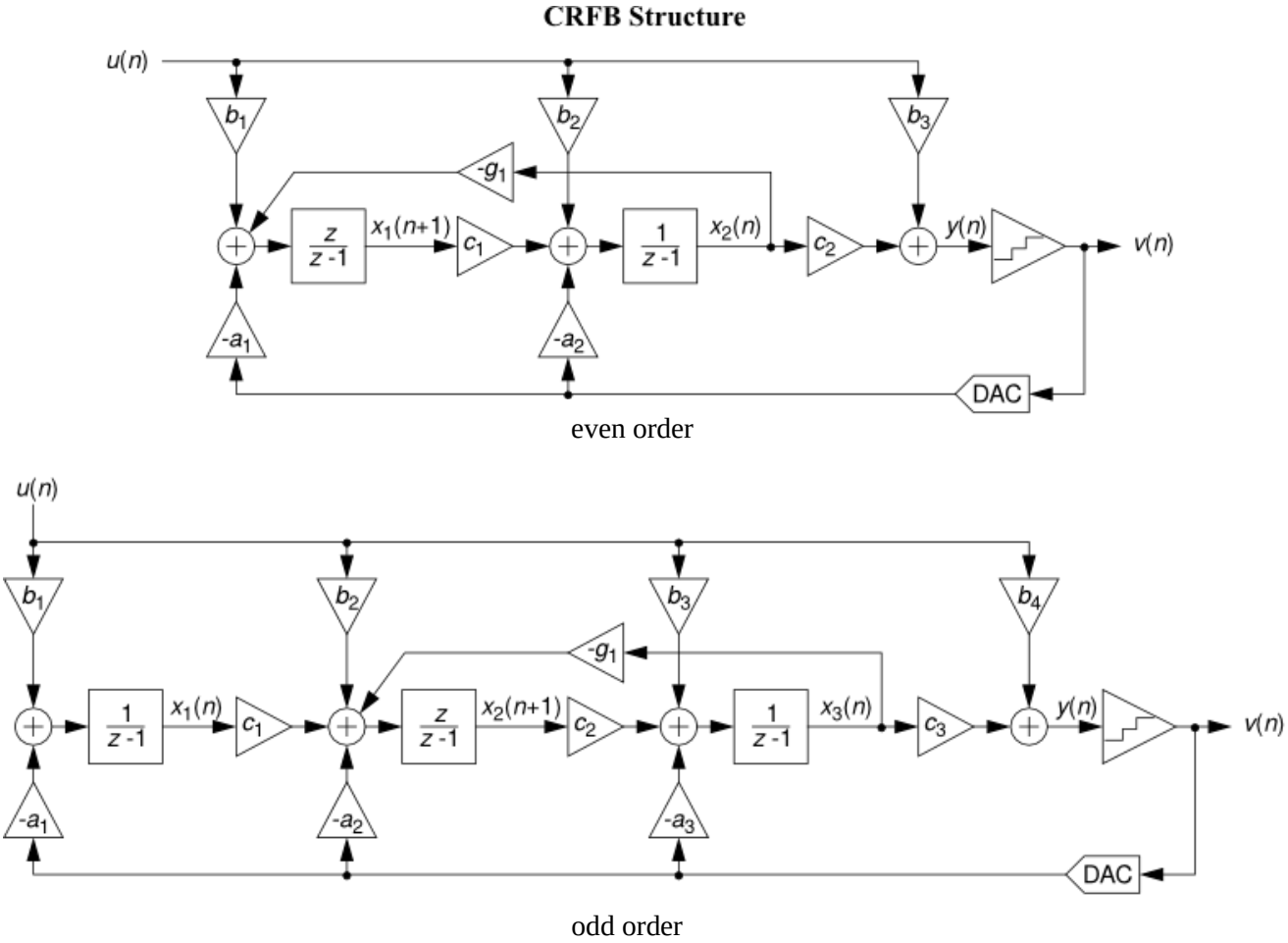


Figure 13 - CIFB ΔΣ modulator architecture example [14]

2 FPGA DESIGN

This chapter discusses the FPGA HDL implementation part of the DAC. Most of the DAC is implemented in the FPGA, except for the 4-element PDM DAC array and the analog output stage, which are realized in dedicated hardware. The architecture has been designed to be able to accept input sample rates of 44.1 kHz or 48 kHz and bit-depth up to 32 bits (compatibility with *USB Audio Class*). Sample rates up to 768 kHz are achievable by changing the reconstruction filter state-machine. The overall DAC architecture is shown in Figure 14.

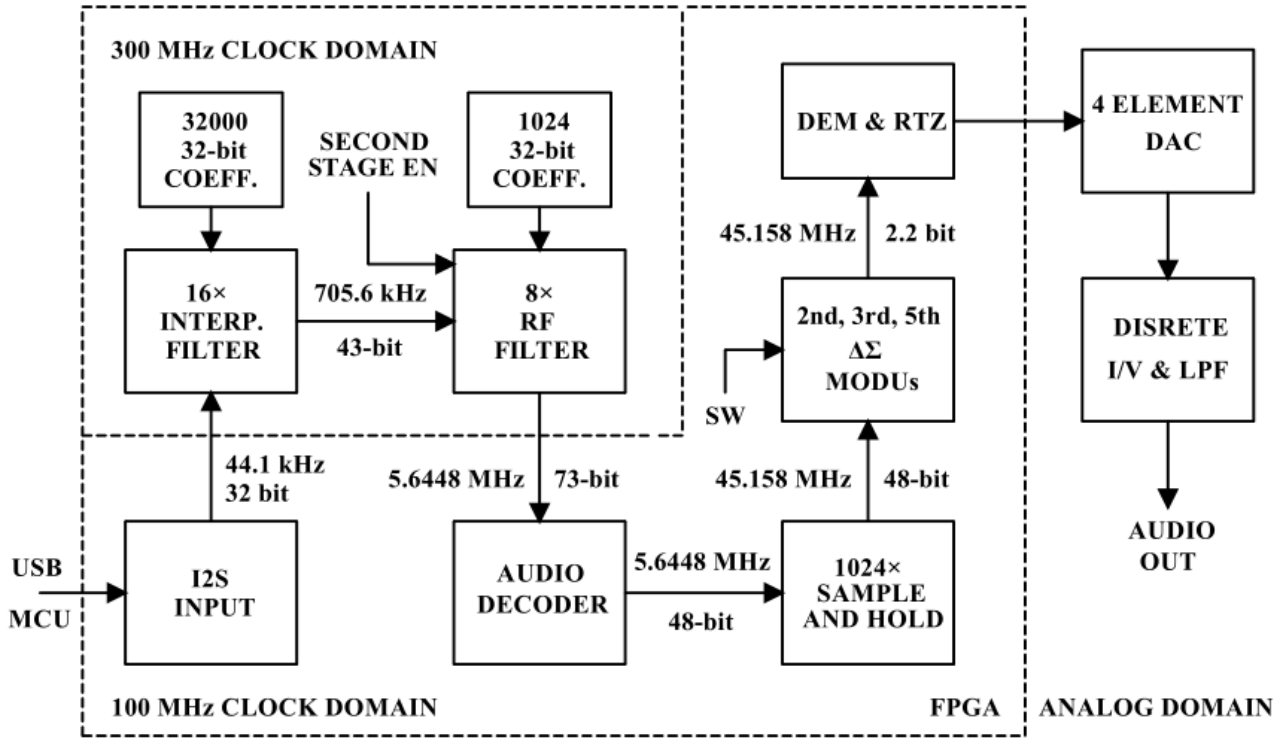


Figure 14 - The ΔΣ D/A converter architecture overview

The USB microcontroller (not shown) communicates with a PC over USB Audio Class 2 standard to obtain the audio signal data. The obtained signal is reconstructed by 16× interpolating *FIR* filter, resulting in a 705.6-kHz (or 768-kHz for 48 kHz input) 43-bit signal. This signal may be optionally further interpolated by a factor of 8 up to 5.6448 MHz. This second interpolation is done to get rid of most of the signal images and their

RF energy. The following audio decoder block extracts the left and right channels and truncates the data down to 48 bits. The *Sample and Hold* circuit completes the interpolation process, resulting in OSR of 1024. The data-stream is then modulated down to a 5-step (2.2-bit) signal by the ΔΣ modulators. The 2.2-bit PDM stream is split into four separate PDM signals, each driving one element of the four-element discrete PDM DAC array.

2.1 Signal interpolation and reconstruction

Upsampling is a process of increasing the signal sample rate by inserting zero value samples between the original samples (*zero-stuffing*). Upsampling creates undesired *spectral images* of the original signal that need to be removed by the filter. The process of upsampling followed by such a filter is called *interpolation*. [8]

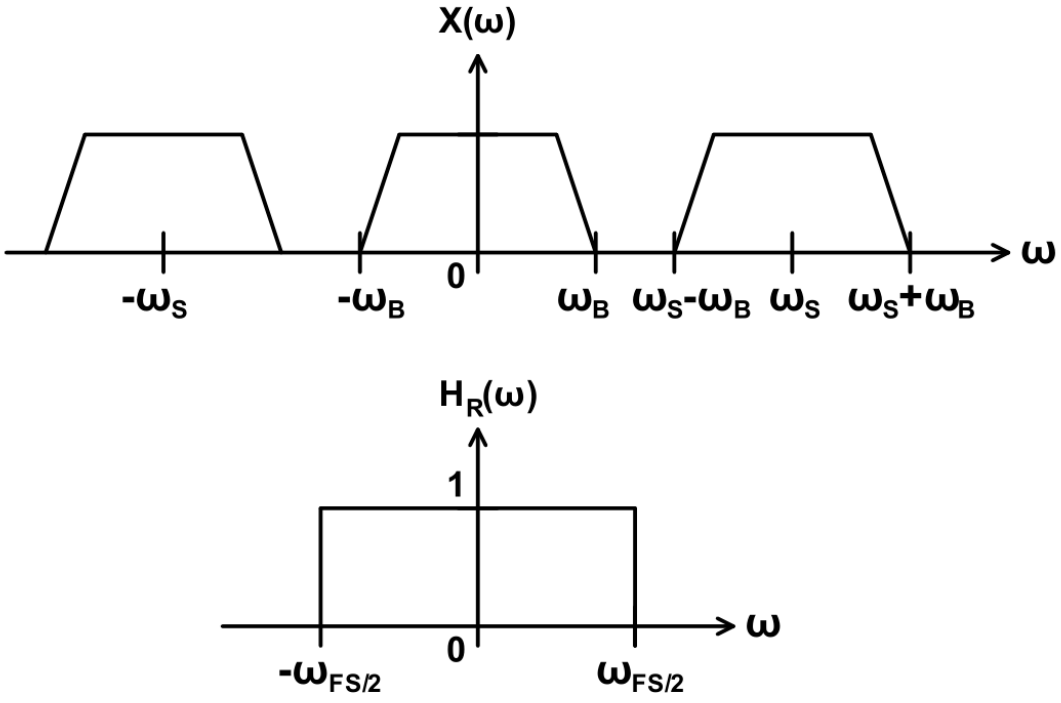


Figure 15 - A) Signal spectrum images around s sampling rate of ω_S

B) Normalized boxcar function response H_R

The upsampling process is illustrated in Figure 15A. A signal with a bandwidth of ω_B is mirrored around the sampling rate multiples of $n\ \omega_S$. These signal images extend back to the bandwidth of $\omega_S - \omega_B$. Aliasing occurs in case the signal ω_B extends over half the sampling rate of $f_s/2$. Images start to overlap the original signal because there is no longer one solution to the signal reconstruction. Therefore, the signal has to be band-limited according to the Nyquist sampling theorem defined as

$$f_B = \frac{f_S}{2}\ , \tag{2.10}$$

where f_B is the signal bandwidth and f_S is the sample rate.

Aliasing might occur when converting a signal from a higher rate to a lower rate. This process is called *decimation* in the discrete-time domain, and it is known as A/D conversion in case of infinite rate (continuous) input signals. Low pass filters are used to limit the effects of aliasing.

Similar is true for interpolation. However, the spectral images shift further from the original signal instead of shifting towards 0. The spectral images are defined

$$\omega_S = \mathrm{L}\ \omega_B, \tag{2.11}$$

where L is the interpolation factor. These images have to be removed to reconstruct the original band-limited signal. Removing them is accomplished by a low-pass filtering at the Nyquist frequency f_B by a *half-band* filter illustrated in Figure 15B as a *normalized boxcar function*.

A particular type of *sinc function* is obtained if the boxcar function is transformed to the time domain with Fourier transform. This function is called unnormalized sinc function and for a half-band filter is defined as

$$\mathrm{X}_R = \mathrm{sinc}\ \frac{\pi t}{T}, \tag{2.12}$$

where T is the sample period, and t is time. Figure 15 shows a representation of a partial unnormalized sinc function. One of the function properties is that it is zero at sample periods T (where the signal sample lays) and non-zero everywhere else.

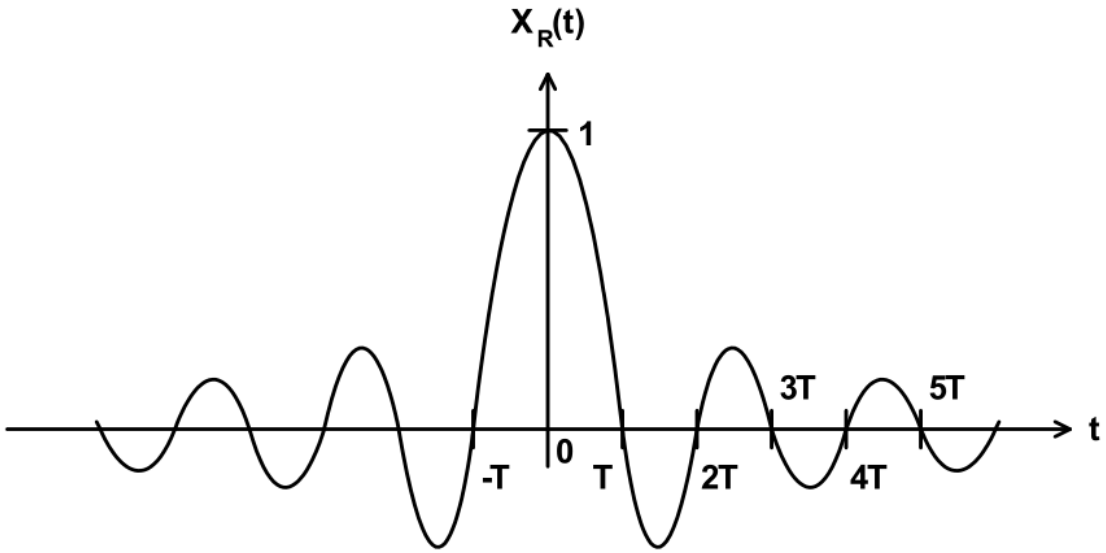


Figure 16 – Part of the unnormalized sinc function

The reconstruction process can be represented as a convolution

$$\mathrm{x(t)} = x_S * x_R, \tag{2.13}$$

x_S is the sampled signal, x_R is the reconstruction filter impulse response, and x is the original band-limited signal. Equation (2.13) can be extended for the discrete-time as

$$\mathrm{x(t)} = \sum_{n=-\infty}^{\infty} x[n] sinc(\frac{\pi}{T}(t - nT)). \tag{2.14}$$

Equation (2.14) signifies that the reconstructed output signal is a sum of an infinite number of sinc functions time-spaced by periods T and multiplied by an input sample x. This sum reconstructs the intermediate values between samples precisely as they were in the original waveform. [8]

Summing an infinite number of sinc functions is impossible in practice. Hence a sinc approximations are made that are usually around 100-500 samples long. [8] However, approximated filters do not reconstruct the time domain information of complex audio signals correctly, despite the fact that their frequency domain performance is sufficient. Reconstruction filters are usually non-causal. Non-causal systems need information about the signal before T = 0, as seen in the impulse response in Figure 16.

The correct sinc function sum can be made by delaying the input signal by half of the filter length. This delay gives the convolution engine the required "future information." The approximated sinc function filters are designed with quick decay of the sinc function and reaching zero about a few hundred samples from the center main-lobe. Hence the filter impulse response differs from ideal sinc functions vastly, resulting in amplitude errors in time when reconstructing the signal. These timing errors are especially significant when transients occur in the recorded music as there is an insufficient number of sinc functions in the past and the future to add up and reconstruct the original signal shape, amplitude and position in time. [7]

Approximated and short reconstruction filters are possibly one of the reasons why ordinary consumer audio DAC chips might sound muddy and undefined. The fact that many people are preferring the *NOS* (Non-Oversampling) DACs is maybe due to the sound signature of these approximated filters. NOS DACs use zero-order interpolation (sample and hold) and rely on the analog audio system (and the human ear) to act as the reconstruction filter. NOS DACs usually sound better than ASIC DACs with musical instruments and voices being more transparent and more true to life. However, they are also burdened with significant group delay errors and aliasing that some people may like.

2.1.1 Reconstruction FIR filter design

According to (2.13), the signal reconstruction is defined as the convolution between the sampled signal and the ideal sinc function. Commonly used and well known *FIR filters* (Finite Impulse Response filters) are time-domain convolution engines

$$y[n] = \sum_{i=0}^{N} b_i \cdot x[n-i], \tag{2.15}$$

where N is the number of coefficients (referred to as taps) of a discretized filter function b_i, x is the input signal, while y is the output. Therefore, they make ideal candidates to implement a sinc reconstruction filter. The ideal sinc reconstruction filters require infinite computation time as they have an infinite coefficient length. Compromises have to be made to make a practical reconstruction filter.

These include

- finite time-resolution – a finite output sample rate has to be chosen when operating in the discrete-time domain,
- finite sinc length – longer sinc impulse response will better reconstruct the signal,
- window function – a window function must be applied to sinc impulse response to correct for the frequency response of a finite coefficient length,
- finite coefficient size – coefficients have to be quantized.

Only a finite output sample rate of the reconstruction filter can be achieved when operating in the discrete-time domain. This finite sample rate can also be thought of as sampling and quantizing the sinc function impulse response at regular intervals. Ideally, the output sample rate of the reconstruction filter would be equivalent to the input rate (or OSR) of the ΔΣ modulator. This filter would be incredibly computationally intensive, even though it can be realized. Compromise has to be made. An industry-wide approach is to reconstruct the signal to a high enough sample rate and apply zero-order interpolation after the reconstruction to achieve final OSR.

To evaluate what sample rate has to be reached for the waveform to be fully reconstructed, let us consider some fundamental limitations of human hearing. A high-frequency response limitation of about 20 kHz at most (for young people) is a well-known fact. On the other hand, studies [9] [10] have shown that the human brain uses time-domain information to evaluate its surroundings. This effect mainly concerns voice and instrument localization (both in width and depth – referred to as soundstage), pitch, and timbre. These factors affect the overall realism of the reconstructed musical recording.

Psychoacoustic experiments have proven that the human hearing can discern interaural time differences down to 5 or 4 μs [9][10]. This difference in time translates to a frequency band of about 200 to 250 kHz. Therefore a reconstruction filter with an output sample rate of 705.6/768 kHz should be sufficient as it has over 350 kHz of the Nyquist bandwidth. Output sample rates are 16× multiple of the input sample rate allowing a fully synchronous design.

Choosing the filter length is a compromise between better waveform reconstruction and computational complexity.

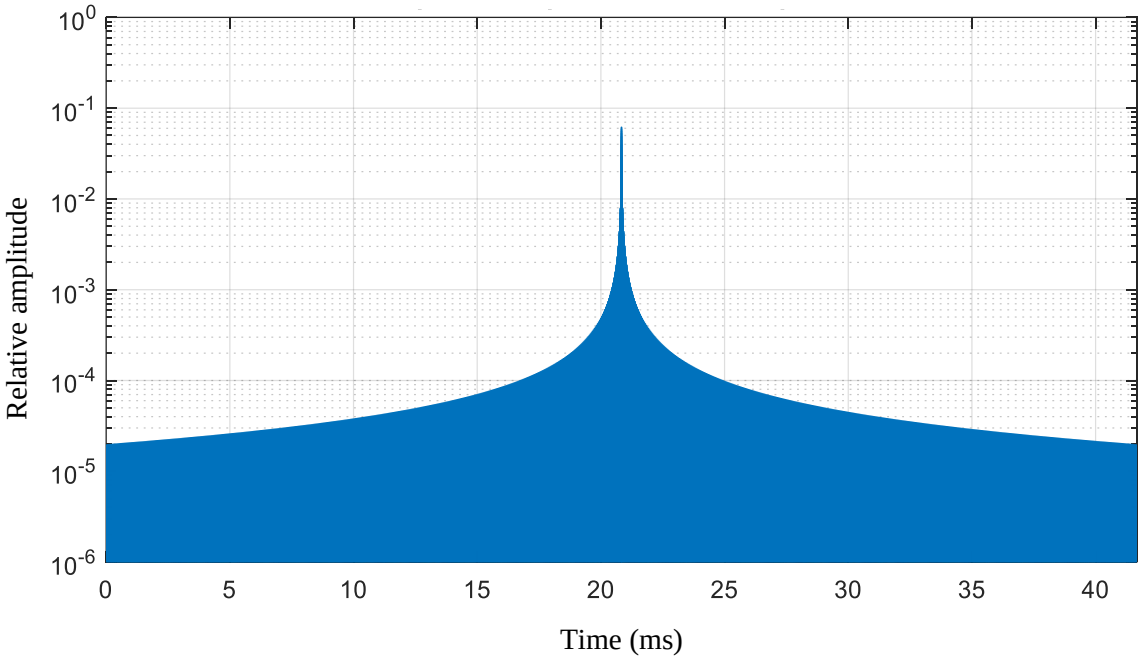


Figure 17 - Envelope of a 32000-point sinc function at 768 kHz sample rate

Figure 17 shows a 32000-coefficient (also referred to as tap) sinc function envelope – an impulse response represented with a logarithmic amplitude axis. It was sampled at 768 kHz, a 16-times the input rate for a 48 kHz input. In 32000 samples, the slowly decaying sinc function only able to decay to

$$1 \cdot 10^{-5} \approx -100 \text{ dB}.$$

While the maximum sinc function value is

$$66.25 \cdot 10^{-2} \approx -24.1\, dB,$$

resulting in a maximum reconstructed dynamic range of about 75 dB. About one-million-tap long sinc response would be needed if 100 dB of the reconstructed dynamic range would be considered sufficient. Hence a 32000-tap function was chosen as a compromise between computational intensity and reconstruction power.

While it is possible to use a limited length sinc impulse response without modifications, the function has not decayed below quantization precision (32-bit in this

case). A sudden cut-off at its edges will occur. This coefficient cut-off will limit the frequency domain performance, as seen in Figure 18.

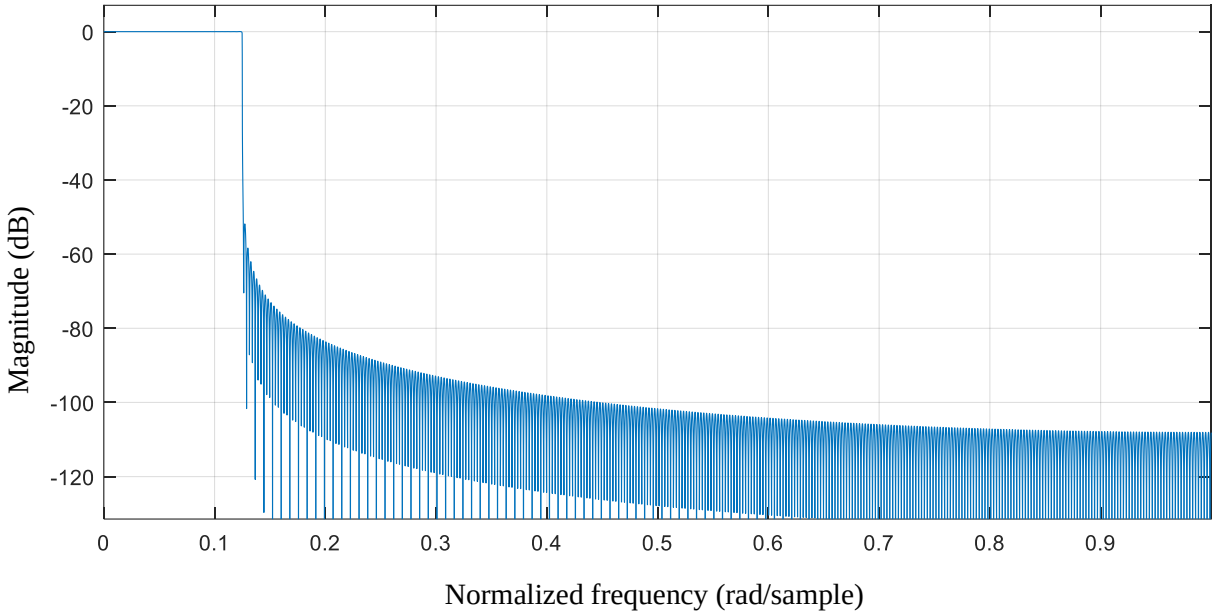


Figure 18 - Frequency response of 32000-tap sinc function not using a window function

A window function can be applied over the sinc impulse response to avoid such behaviour in the frequency domain. A window function will gradually bring the sinc function to zero at the edges. Most popular windows used to design FIR filters, for example Hamming, Kaiser or Chebyshev, change most of the sinc function coefficients to allow for better spectral response, hence they are unsuitable to be used for reconstruction filter design.

The absence of a suitable window function leads to the need to create a custom window function. The proposed function can be seen in Figure 19. Three types of windows are overlayed in this graph – a no window (Rectangular), Chebyshev window (best frequency response), and a custom window designed as a compromise between time-domain reconstruction and frequency-domain performance.

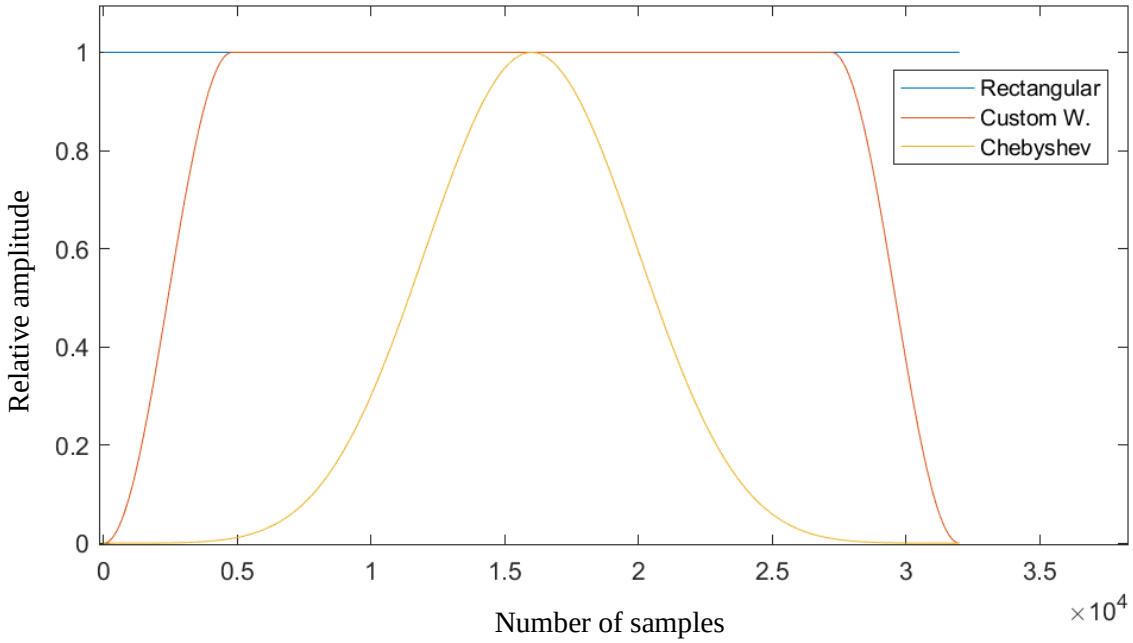


Figure 19 – The window functions

The custom window is defined as

$$w(x) = \begin{cases} \frac{1}{2}\,\mathrm{x}\,\sin\left(n\pi + \frac{\pi}{2}\right) + \frac{1}{2}; n \leq Ls \\ x; \qquad\qquad Ls < n < \frac{L}{s} \\ \frac{1}{2}\,\mathrm{x}\,\sin\left(n\pi - \frac{\pi}{2}\right) + \frac{1}{2};\ n \geq \frac{L}{s} \end{cases}, \tag{2.16}$$

where L is the number of samples, x is input vector of length L in range $\{0,1\}$, n is the sample position, and s is an attenuation ratio describing how steeply is the input vector attenuated. Appling this window with $s = 0.15$ yields a frequency response in Figure 20.

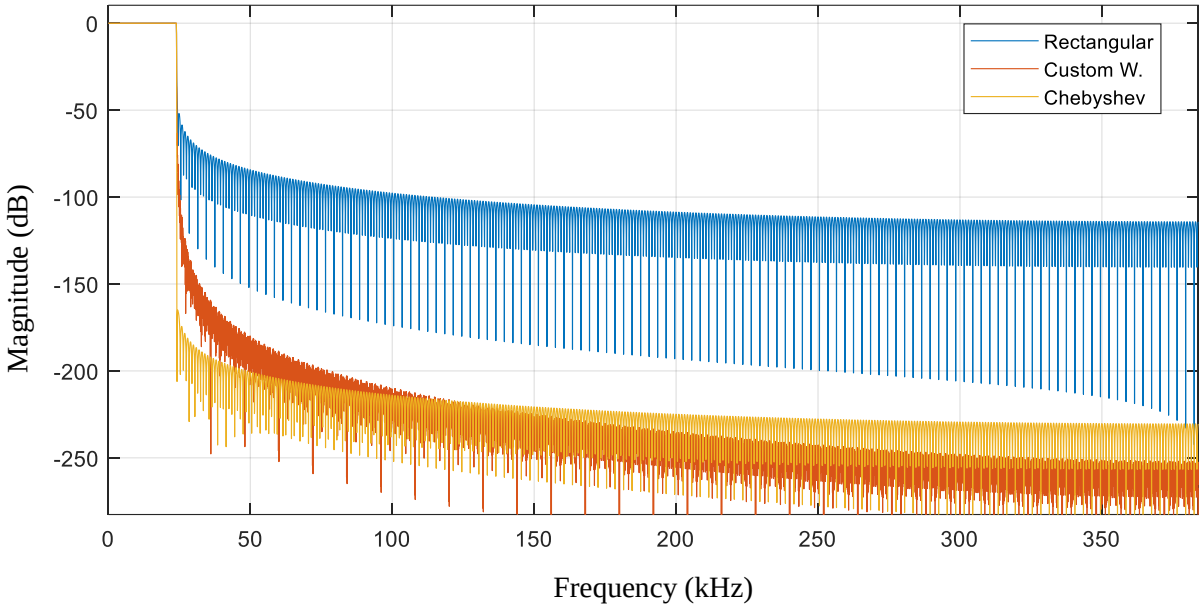


Figure 20- Frequency magnitude response of filters designed with windows from Figure 19

While the Chebyshev window has a superior frequency domain response, it is not able to reconstruct the signal to the same degree as a filter designed with the custom window can. That situation can be observed more closely as a time-domain sinc envelope shown in Figure 21.

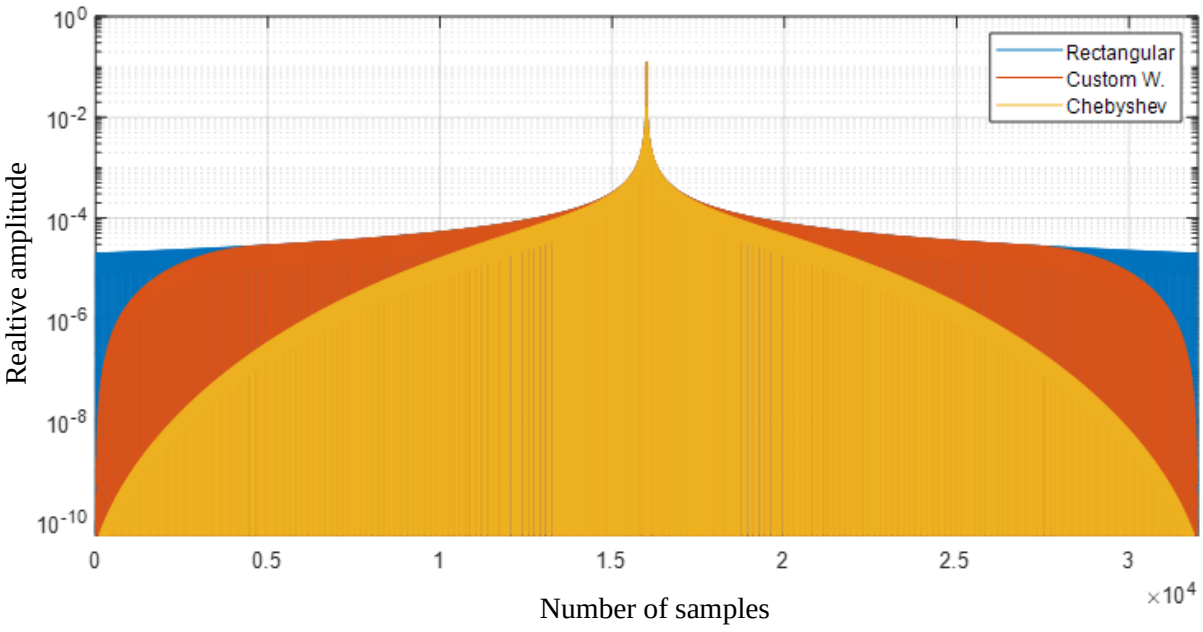


Figure 21 - Envelopes of a 32000-point sinc functions windowed with windows from Figure 19

Chebyshev windowed filter from Figure 21 has the best frequency-domain performance, but decays very quickly. On the other hand, when the sinc function filter is designed with the custom window, it follows the ideal impulse response closely up to a point defined as $s = 0.15$. The sinc response starts to decay quickly to zero at 0.75 multiple of L/2 distance from the center main-lobe.

Implementing this reconstruction filter requires finite precision coefficients. Hence the filter was imported into MATLAB FilterDesigner tool and quantized to fixed-point 32-bit Q format, of which 31 bits are fractional. The finalized reconstruction filter frequency response is shown in Figure 22. Table 2 is a filter properties summary.

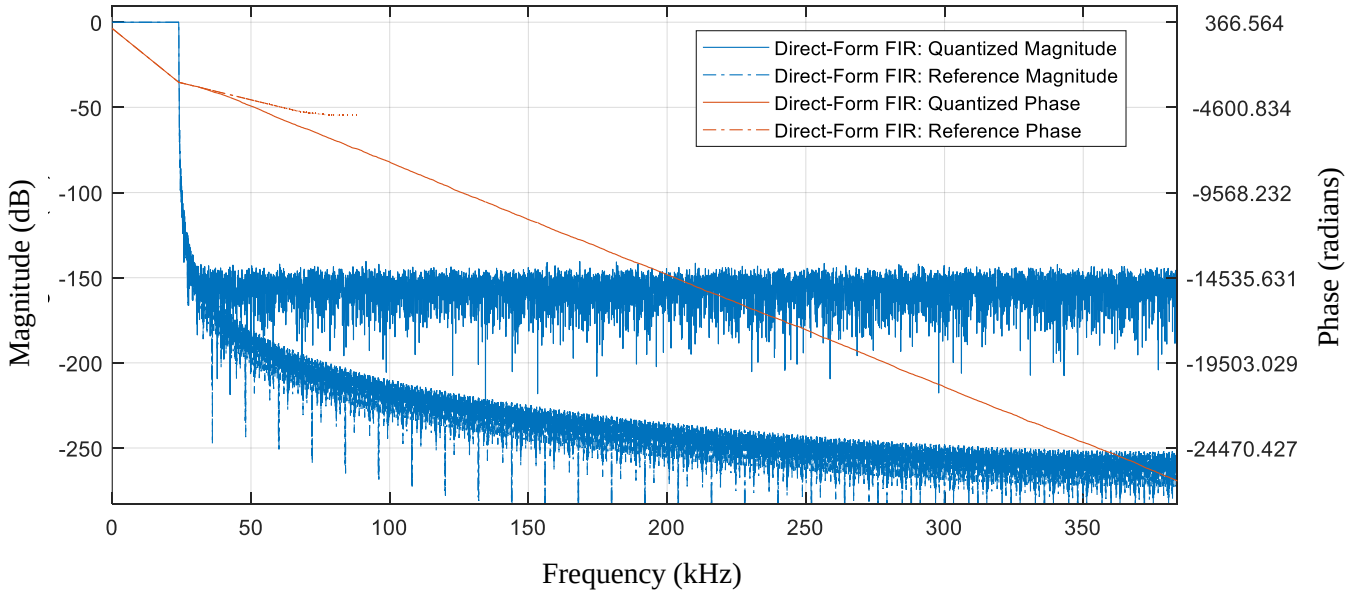


Figure 22 - Magnitude and phase frequency response of the finalized reconstruction filter

Table 2 - Designed interpolator filter parameters, fs = 48 kHz

Parameter	Value
Passband (+- 0.01 dB)	0 Hz - 24 kHz
Stopband	28 kHz – 384 kHz
Stopband attenuation	< -145 dB
Frequency at -100dB point	24.5 kHz
Frequency at -3 dB point	24 kHz
Passband ripple	< 0.1 dB

2.1.2 RF FIR and zero-order filter design

The 16× reconstruction filter is followed by another 8× interpolating FIR block in Figure 14. This second-stage FIR interpolator is significantly shorter, but runs at a much faster rate. Its primary purpose is to reject RF images that would be created if the zero-order hold block followed the reconstruction filter directly. Figure 23 illustrates the situation without RF filter.

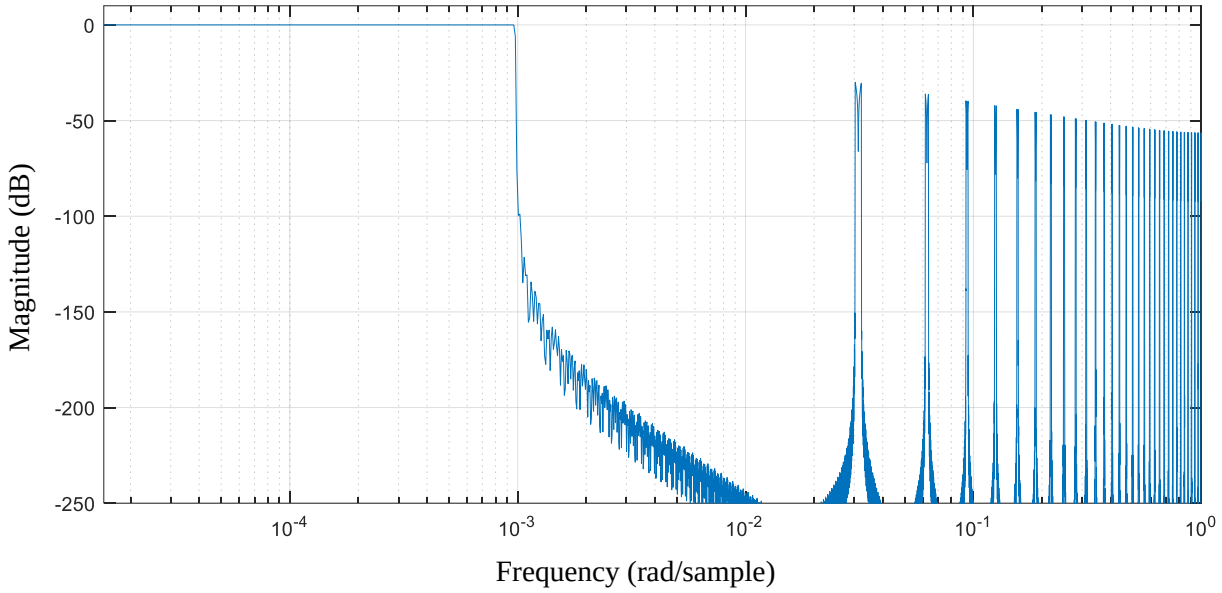


Figure 24 - Reconstruction 16× FIR filter followed by 128× ZOH

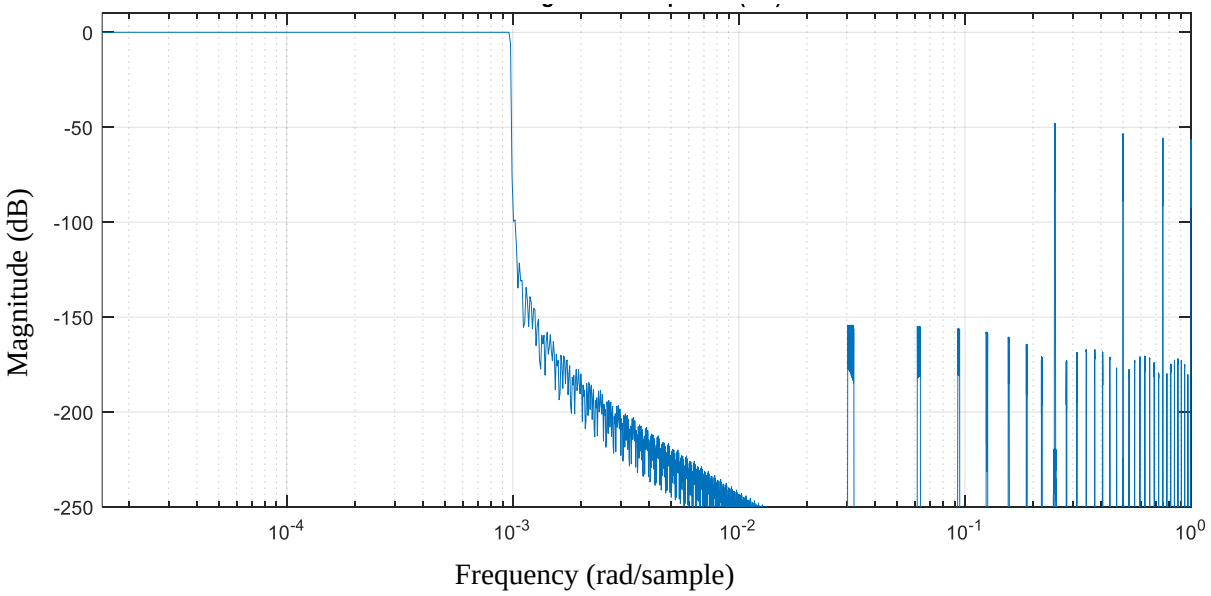


Figure 23 - Reconstruction 16× FIR filter followed by 8× RF FIR filter and 8× ZOH

Figure 24, on the other hand, shows a situation where the RF FIR filter is employed. This RF filter rejects the spectrum images up to 128 OSR, or about 6 MHz. This RF image rejection might help with modulator stability and possibly eases up the requirements on the analog filter. However, the analog filter requirements are mainly determined by the ΔΣ modulator noise shaping slope. This second stage RF filter can be bypassed to allow for an evaluation of the RF filtering possible benefits.

The RF filter was designed with a Chebyshev window function as the signal is already reconstructed. Chebyshev window is relaxing the tap number requirement to 1024 coefficients. Low tap count is required because the filter is running at relatively high sample rates compared to the 6144 OSR filter block master clock (about 300 MHz). The magnitude and phase frequency response of the RF filter is shown in Figure 25.

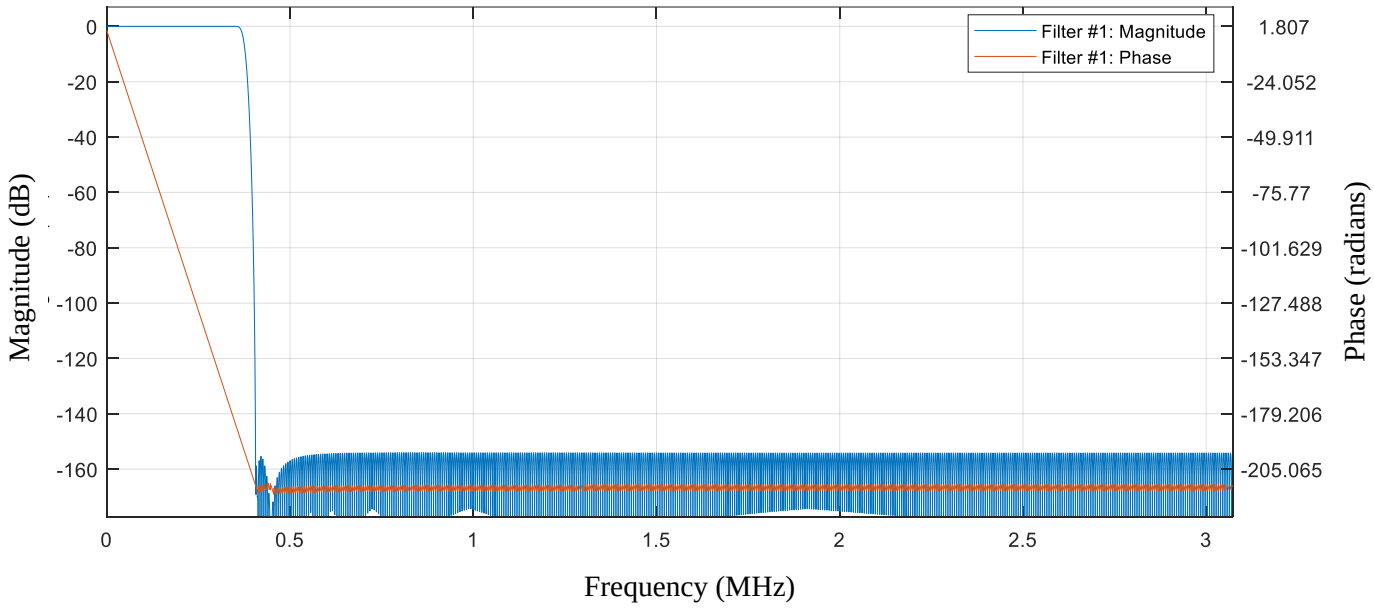


Figure 25 - RF FIR filter magnitude and phase frequency response

Table 3 summarizes the designed RF filter properties. Stopband attenuation is an essential property for this type of filter. The designed filter has a stopband attenuation better than 150 dB – limited by the coefficient quantization.

Table 3 - RF FIR filter properties summary

Parameter	Value
Passband (+- 0.01 dB)	0 Hz - 345 kHz
Stopband	390 kHz – 6144 kHz
Stopband attenuation	< -150 dB
Frequency at -100dB point	384 kHz
Frequency at -3 dB point	357 kHz
Passband ripple	< 0.001 dB

A Zero-order interpolator is the last interpolating block. ZOH is realized by the Sample and Hold block. SH block holds the sample for its oversampling ratio, converting the pulse train samples to a staircase-type waveform in the process. The sample and hold creates a first-order sinc filter magnitude response (not to be confused with sinc interpolation). The magnitude frequency response of 16 OSR (which holds the sample for 16 discrete periods) is shown in Figure 26.

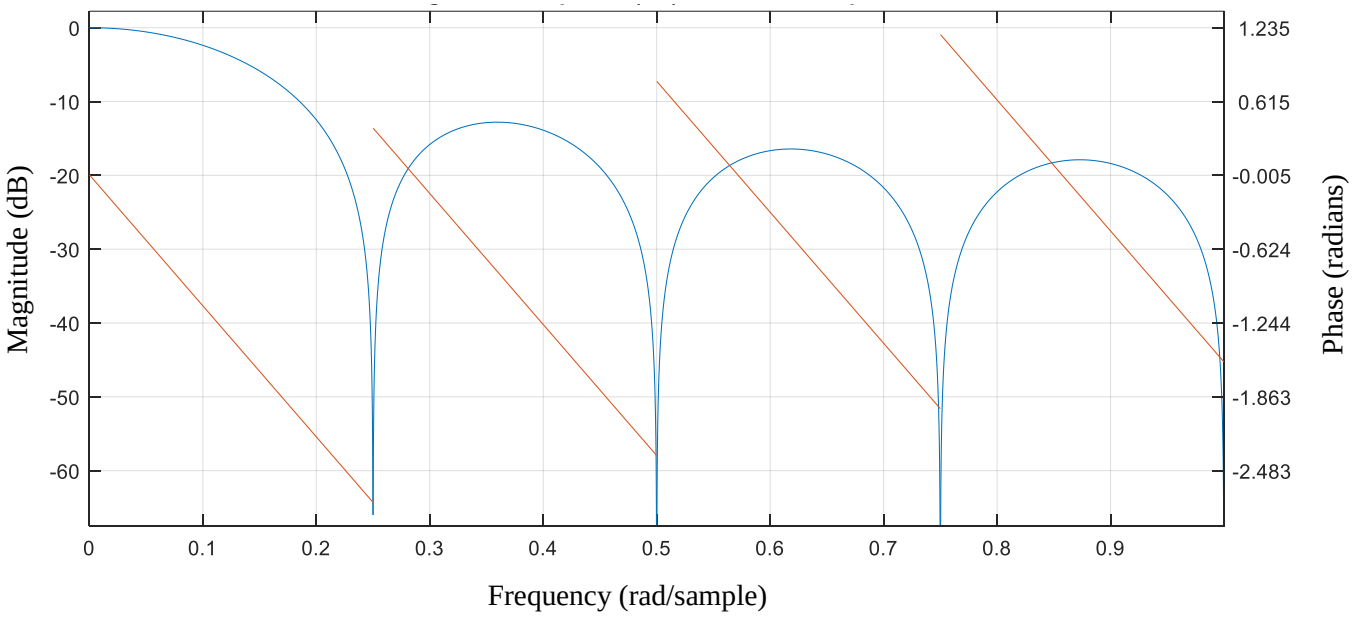


Figure 26 - Zero-order 16× block creating a first-order sinc magnitude frequency response

2.1.3 FPGA implementation

The reconstruction filter is much more computationally intensive compared to audio interpolation filters that are usually designed. Nonetheless, modern FPGAs have enough resources to accommodate even larger filters, as discussed in [7]. Besides the modulator implementation, this fact is one of the reasons why FPGA was chosen over DSP, which do not have nearly enough processing capability. As mentioned, interpolating is a process of upsampling followed by a filter. Since designed filters are types of interpolation filters with 16 (reconstruction filter) and 8 (RF filter) zero-stuffing upsampling blocks, it would be inefficient to spend most of the processing time on multiplication with zero [11].

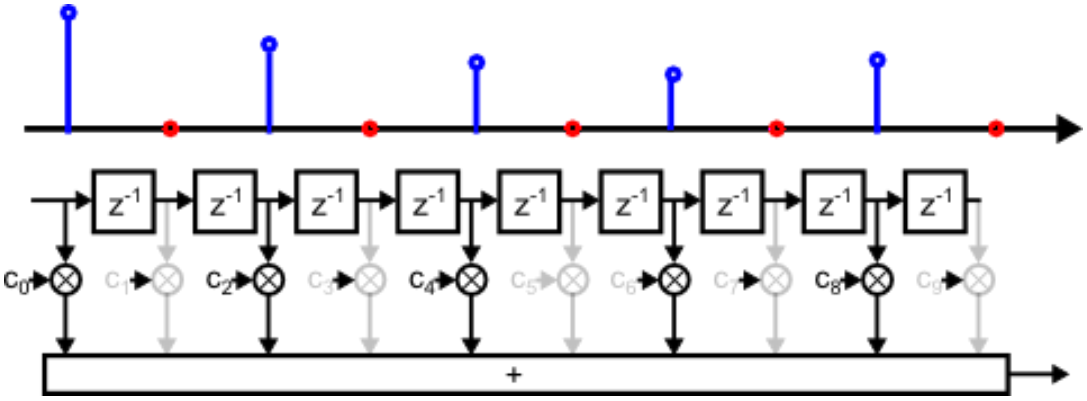


Figure 27 - Example of an 8-tap two-times interpolating FIR filter [11]

Figure 27 is an example of a two times interpolating FIR filter. Upsampled data shifted into the filter are zero at every second sample. Hence only four multipliers are active because the multiply-accumulate operation (*MAC*) is zero at these samples.

When another sample is being processed, as shown in Figure 28, the multipliers that processed zeros are now active. Nevertheless, the other four multipliers are processing zeros again, thus the total number of multiplication per cycle has not changed. Only the coefficients have changed. For the 32000-tap 16× reconstruction filter, this would mean that 512000 of MAC operations would have to be calculated per every cycle, but only 32000 of them would have non-zero output.

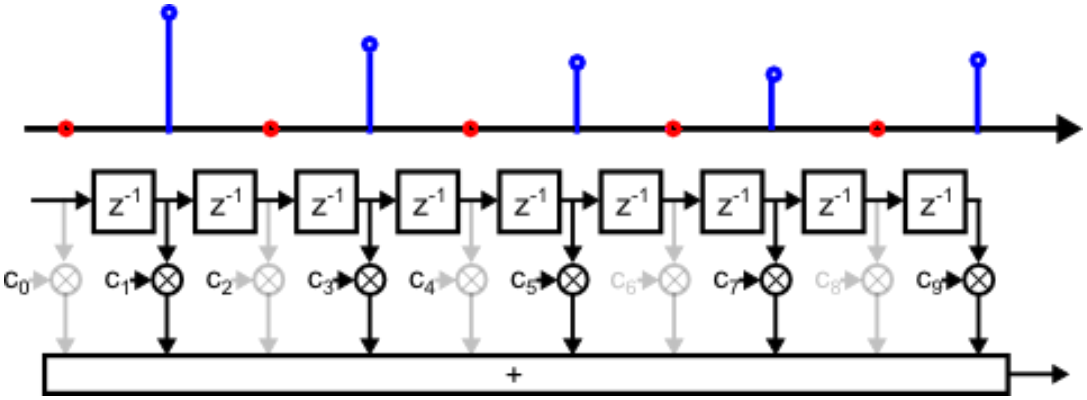


Figure 28 - The interpolation filter from Figure 27 processing the next sample [11]

So-called *polyphase* filters share one multiplier for all OSR operations, while simultaneously multiplexing the multiplier coefficients. Nearly the same operation is done as a result with less work compared to the *direct-form* filter mentioned above (the convolution result is identical in theory, but quantization noise might be different for polyphase filters). An example of the polyphase filter architecture is illustrated in Figure 29.

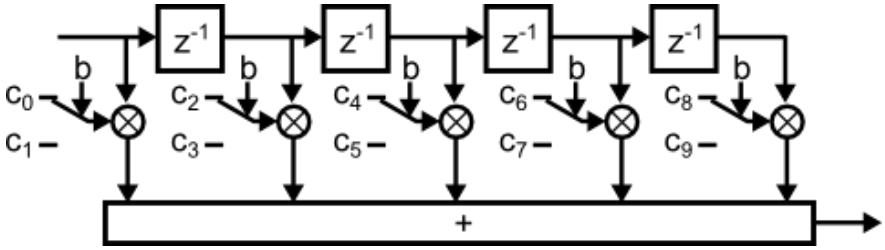


Figure 29 - Architecture of a 2× interpolating polyphase FIR filter [11]

Both reconstruction and RF FIR filters were implemented as polyphase filters with MAC blocks running at 6144-multiple of input rate (294.912 MHz for 48 kHz input). With MAC blocks running at a significantly higher rate than the filter input, they can extend beyond polyphase coefficient multiplexing, as one MAC is able to calculate 6144 outputs per input sample. Consequently, one MAC DSP block is multiplexing between many coefficients resulting in the workload serialization. The coefficients are stored in the FPGA block RAMs and called-out out by the state machine control logic. Block RAMs are also used as data path delay line. One filter is interleaved for both audio channels.

Both FIR filters were compiled into HDL using Xilinx's FIR Compiler 7.2. Resource utilization of the synthesized filters is summarized in Tables 4 and 5. An Artix-7 XC7A35T FPGA is the implementation target.

Table 4 - Reconstruction FIR filter resource utilization

Resource	Utilization
LUT6	280/20800
FF	1412/41600
BRAM	42/50
DSP48 block	51/90

Table 5 - RF FIR filter resource utilization

Resource	Utilization
LUT6	808/20800
FF	886/41600
BRAM	3/50
DSP48 block	19/90

The DSP48 blocks consist of pre and post adders, a combinational multiplier and a simple ALU. The DSP48 blocks are placed in columns with dedicated carry data paths between the blocks. Dedicated pipelining registers are also included in the DSP48 blocks. Each DSP48 block MAC is 25×18-bit input with 45-bit wide outputs (sign-extended to 48-bit). The FIR Compiler combines many of the DSP48 blocks for parallel operation to implement the FIR filter with 32-bit input and 32-bit coefficients.

Most of the sinc impulse response is below the 25-bit or even 18-bit precision of one DSP48 block. However, the FIR Compiler presumes full-scale input coefficients, and hence the bit-depts inside the filter reach up to 72-bits. Optimizations can be done when using just a single DSP48 25×18-bit multiplier for most coefficients after the sinc amplitude has decayed below this precision.

Block RAMs are used for the coefficient storage as well as audio sample delay-line. The BRAM columns are next to the DSP48 columns for convenient interconnections. A lot of Block RAM can be freed-up by scaling the coefficients to 18 or 25 bits after the sinc impulse has decayed below this precision, as mentioned above. Its also possible to use LUT6 configured as RAM for the audio data delay-line, resulting in more freed BlockRAM.

All of these methods, however, require writing a customized convolution architecture and are beyond the scope of this thesis. Symmetry of the filter is exploited for the 1024-tap RF FIR filter, using the DSP48 blocks pre-adders. It is not implemented for the first stage reconstruction FIR filter to allow experiments with non-linear phase filters.

2.2 ΔΣ Modulators

A second-order, third-order and fifth-order ΔΣ modulators were designed. All of them classify as selectable CIFB/CRFB single-loop architecture.

2.2.1 Simulation Testbench

A MATLAB Simulink environment is used to design and simulate these modulators. Figure 30 shows a testbench with stimulus generators for the modulators. Figure 31 illustrates how one modulator block is connected on the testbench.

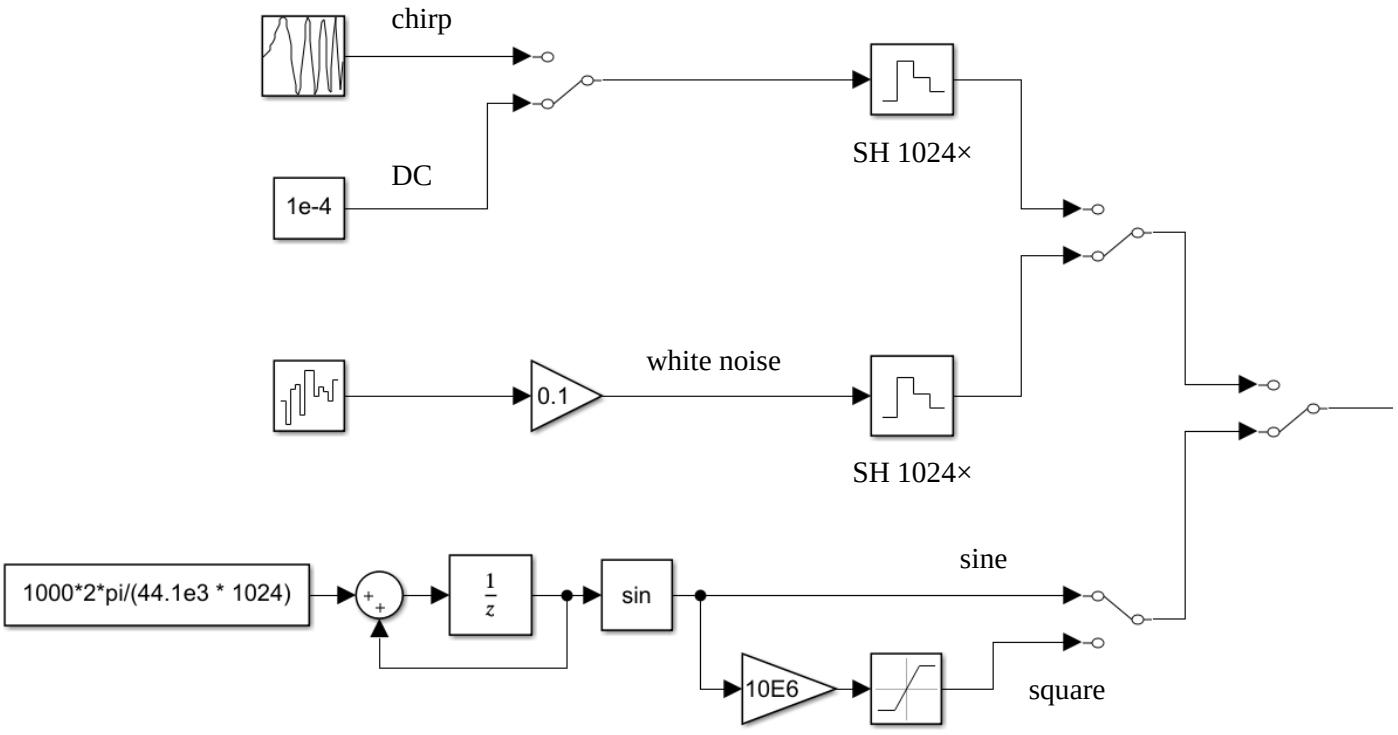


Figure 30 - Stimulus generators in the created MATLAB Simulink testbench

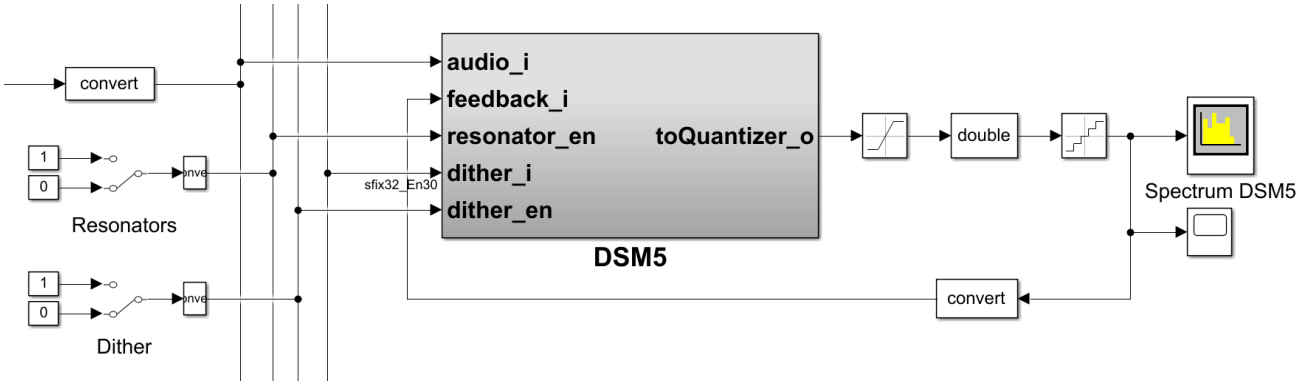


Figure 31 - Example of ΔΣ modulator connection on the testbench

Each ΔΣ modulator is a separate sub-block to allow for direct HDL synthesis into VHDL code with a MATLAB HDL Coder. All of these blocks have an external quantizer feedback loop (which is implemented in the VDHL modulator wrappers), and they are connected to the same stimulus bus. That allows for an easy comparison between modulators at the same conditions. The modulator quantizer block output is a 5-step PDM signal, as seen in Figure 32.

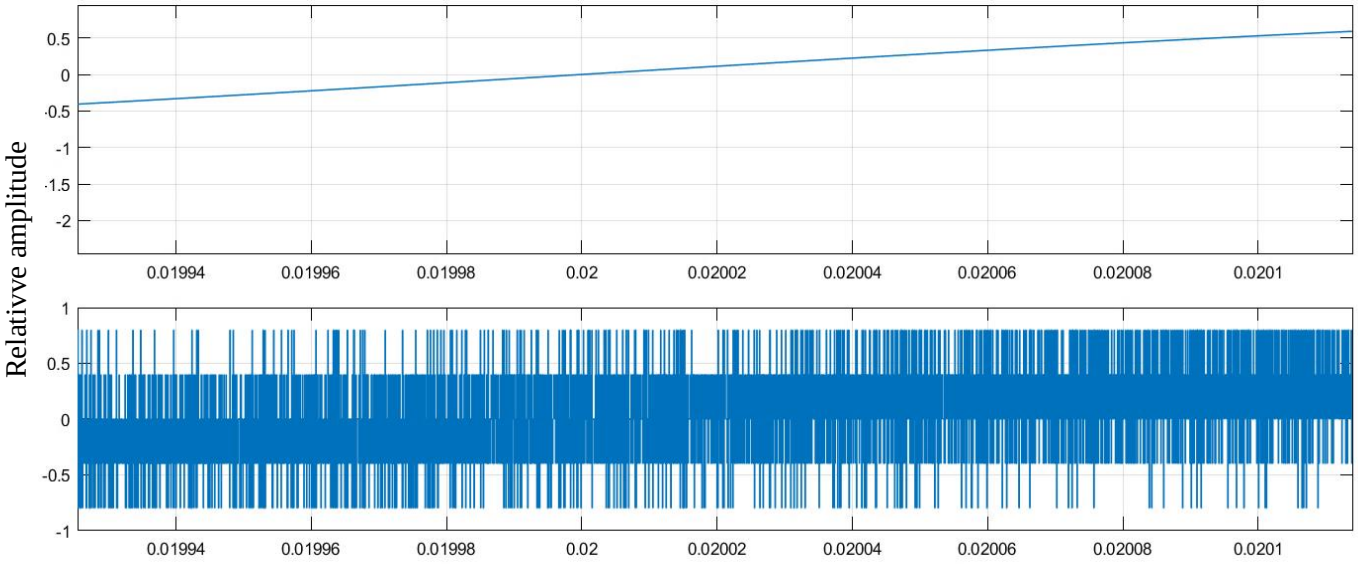


Figure 32 - Multibit PDM output stream of the 5-step quantizer

2.2.2 Modulator Architecture

The ΔΣ modulators share the same selectable CIFB or CRFB loop architecture. The second-order modulator Simulink implementation is shown in Figure 33. The second-order architecture is the simplest implementation and it was chosen for illustration purposes. The third-order architecture has another set of *A*, *B* coefficients, an adder, and an integrator. The fifth-order architecture also has another resonator loop R. Architectures for all modulators are attached at the end of this thesis.

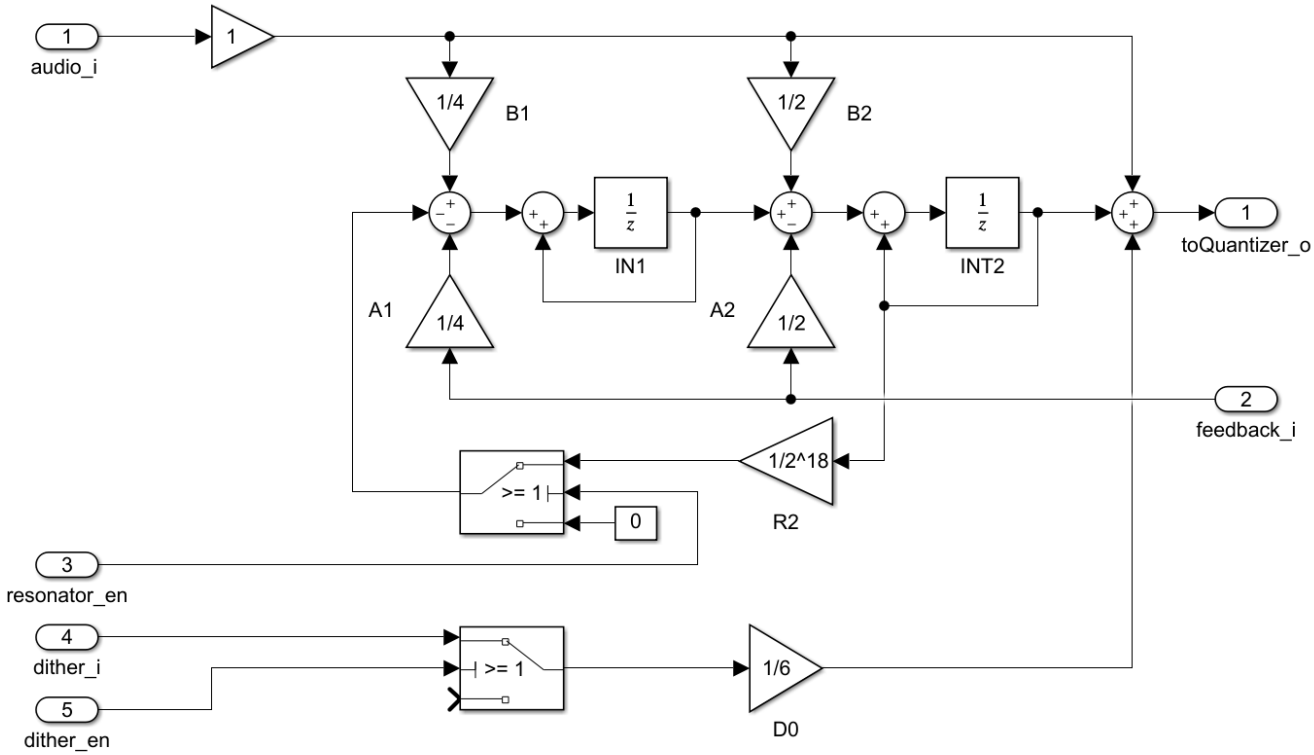


Figure 33 - Simulink model for second-order ΔΣ modulator block

Each elemental block comprises of a delaying integrator (INT), an add-subtract block and feedforward (*B*) and feedback (*A*) coefficients. The output adder block adds dither before quantization, and the feedforward path increases stability.

The resonator loop R creates zero at a specific frequency, which may consequently increase the inband SNR. Selecting between the Cascade of Integrators, Feedback (CIFB) or the Cascade of Resonators, Feedback (CRFB) architecture is done by disabling or enabling this resonator loop, respectively. The fifth-order modulator has two of these resonators, allowing it to create two zeros in the spectrum.

All the coefficient gain blocks are implemented as a power of two. No resources are then used for the coefficient gain blocks because simple bit-shifts are used for the hardware implementation. The modulator is implemented in a 48-bit fixed-point Q-format. Nearly all modulator sections are implemented as Q2.46 format, but the integrators are implemented as Q3.45 format to allow for more overload headroom.

The choice of the fixed-point Q-format over floating-point increases the simulation time significantly on any modern CPUs with AVX instruction set extension. However, the HDL coder can compile the fixed-point Simulink blocks into VHDL code, allowing for very convenient modulator prototyping.

2.2.3 Simulation Results

The testing was done at a 44.1 kHz input rate with a modulator OSR of 1024. Chebyshev window function with a sidelobe attenuation of 600 dB for the FFT. The RBW of the FFT is 20.32 Hz.

The noise-shaping performance of all three modulators in CIFB mode is shown in Figure 34. The noise shaping-slopes are typical for second, third and fifth-order ΔΣ modulators with most zeros at DC. Poles are placed just inside the complex *z* unit circle, resulting in aggressive noise shaping. Some zeros are moved from DC near the Nyquist rate to ensure stability.

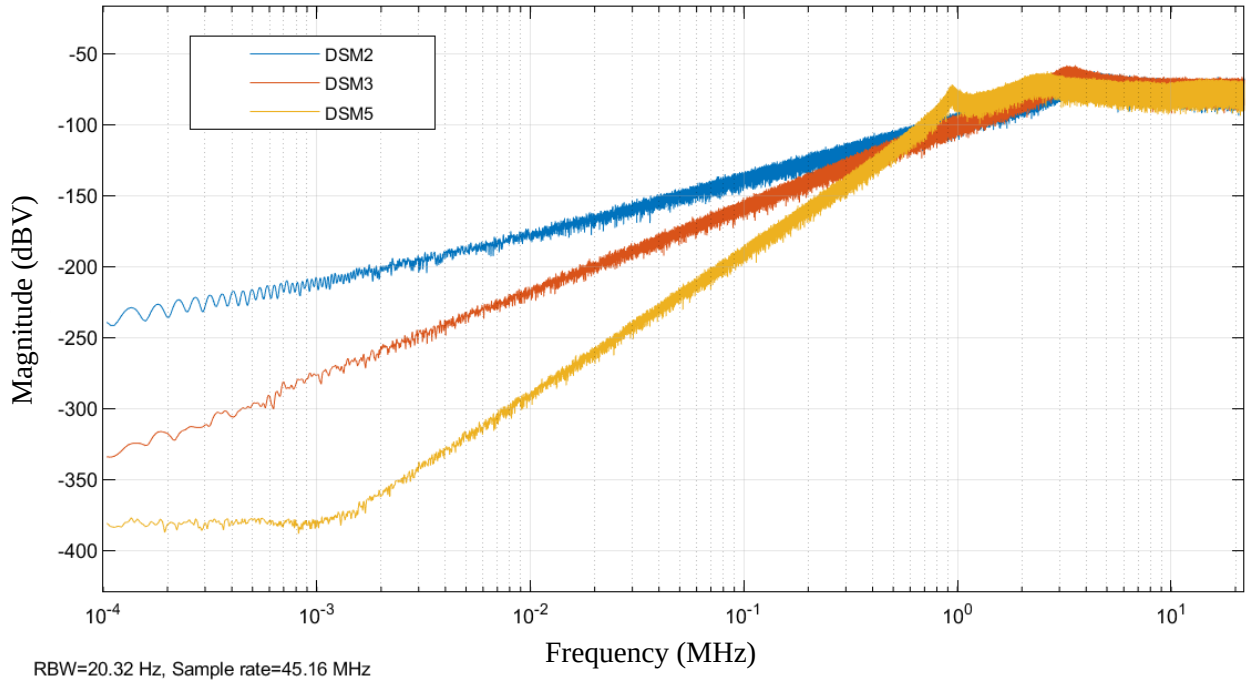


Figure 34 - Output spectrum of the CIFB ΔΣ modulators with no input signal

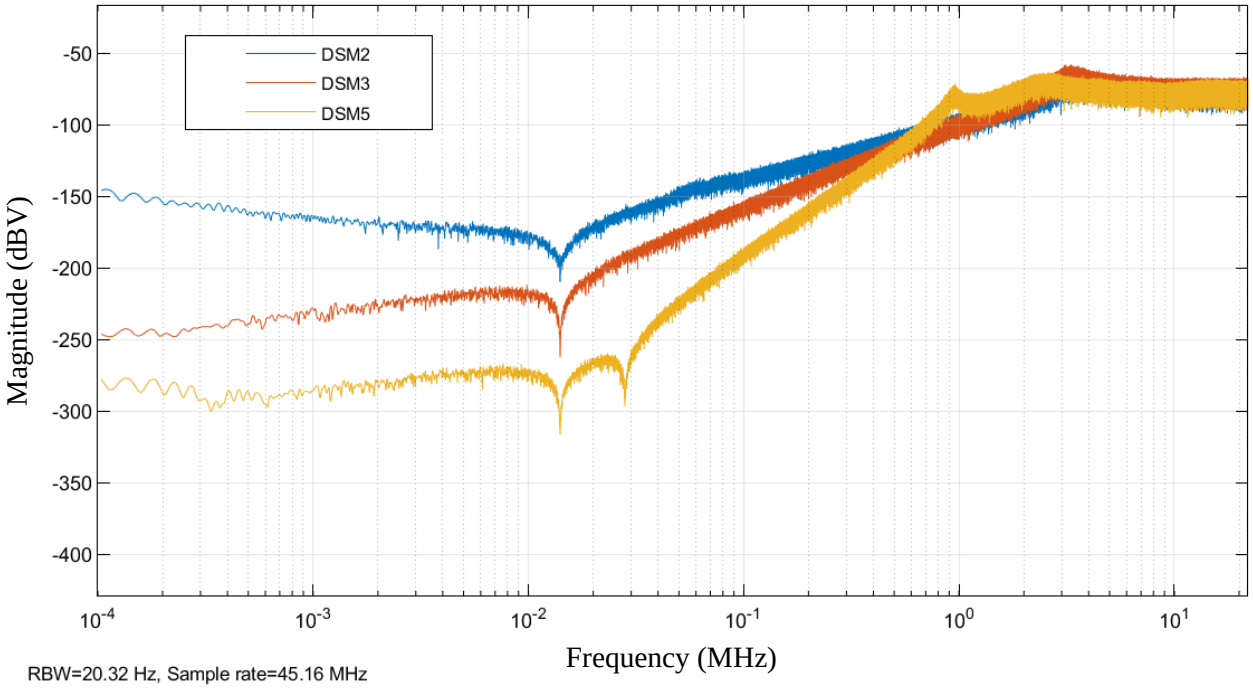


Figure 35 - Output spectrum of the CRFB ΔΣ modulators with no input signal

The modulators are switched to CRFB mode when the resonators are enabled. The resulting noise-shaping slopes are shown in Figure 35. The resonators shift some zeros from DC to the audio band extending the dynamic range in the process.

The spectrums in Figure 36 are the CIFB loop modulators when a -1 dBFS input 1 kHz signal is applied. Second and third-order modulators exhibit little signal harmonics at 5 kHz and 6 kHz. The fifth-order modulator does not have these harmonics, but it exhibits some harmonic content on the noise-shaper slope.

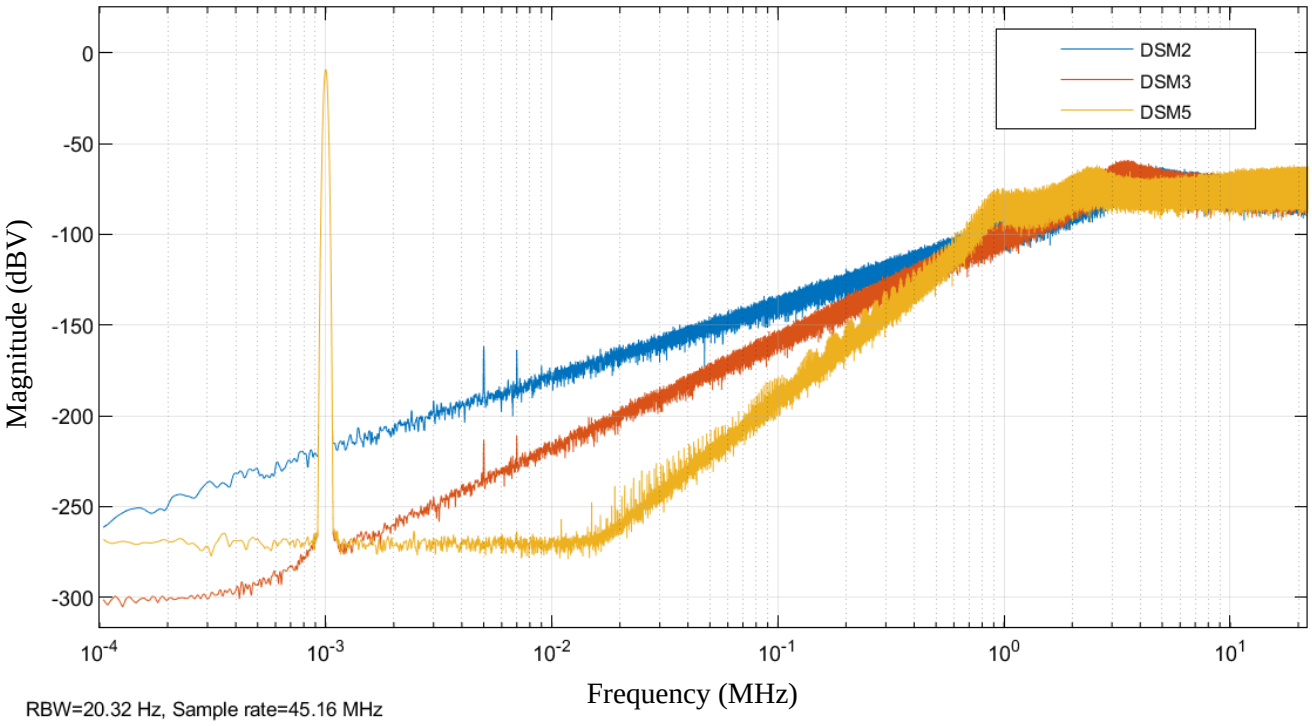


Figure 36 - CIFB ΔΣ modulators with 1-kHz -1 dBFS input signal

Even the fifth-order modulator starts to exhibit the 5 kHz and 6 kHz harmonics when the modulators are switched to CRFB loop mode, as can be seen in Figure 37.

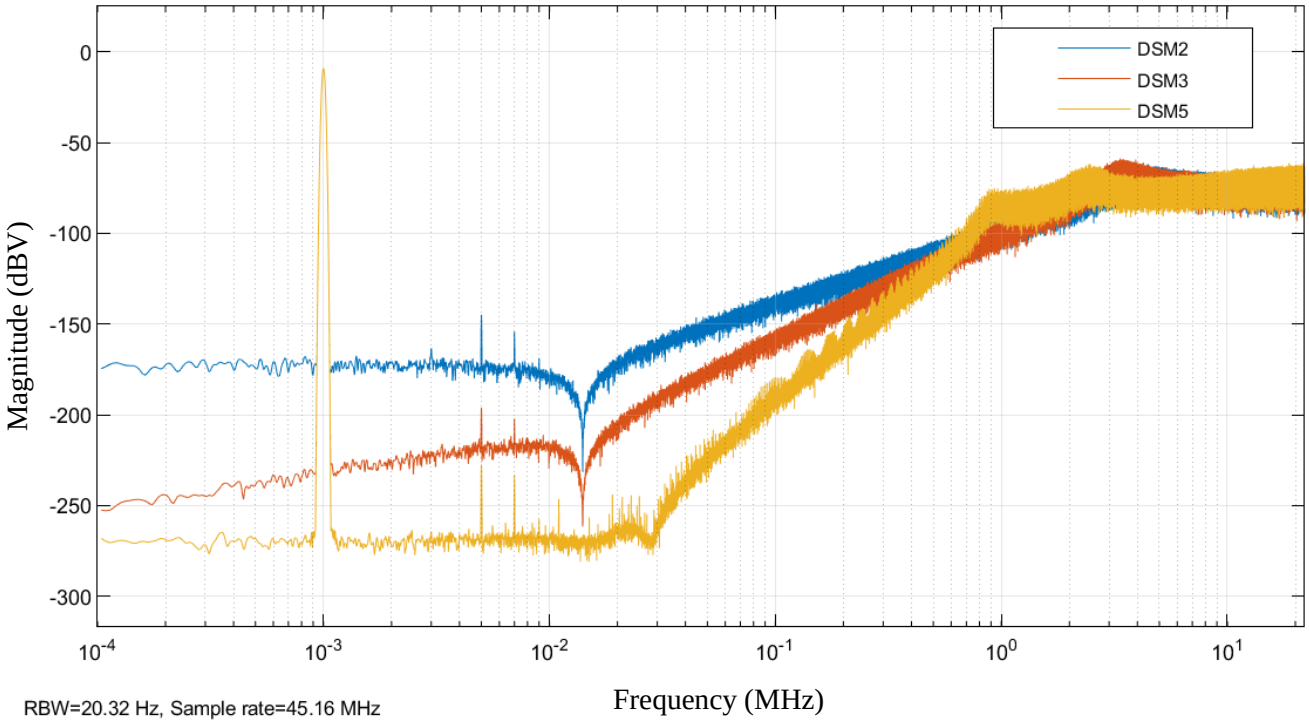


Figure 37 - CRFB ΔΣ modulators with 1-kHz -1 dBFS input signal

Insufficient quantizer dithering might explain such behaviour. According to paper [12], it would take ±1 LSB of TPDF dither to linearize the quantizer properly. While this is possible for the 5-step (5 LSB) multi-bit quantizer used in this thesis, it would consume a large portion of the quantizer resolution. Perhaps using PWM instead of PDM would be a better solution. PWM would allow for more LSBs to be encoded, but would lower the OSR of the ΔΣ modulators significantly.

The fifth-order CIFB modulators flat noise floor is caused by the Q3.45 fractional part format limiting the first integrator, i.e. rounding errors. This limitation is not the case for second and third-order modulators as they do not have this block.

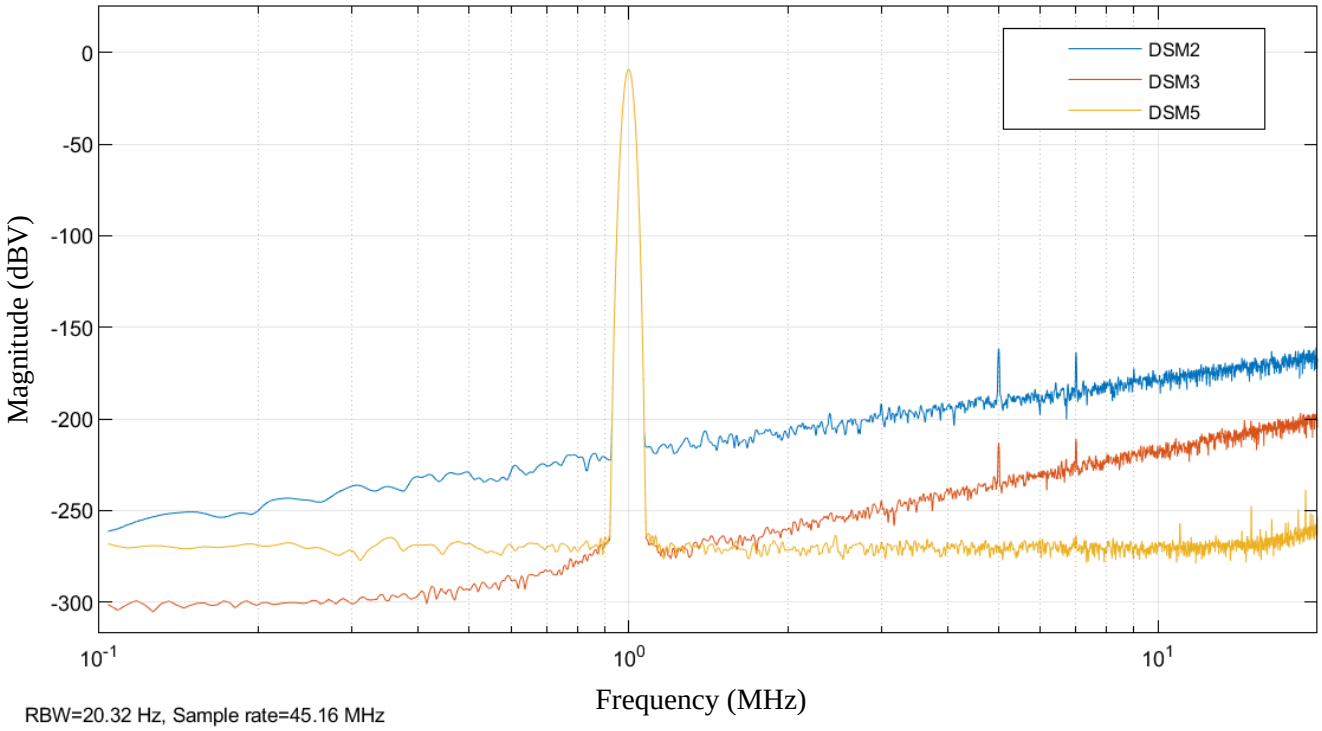


Figure 38 - In-band spectrum of CIFB ΔΣ modulators with 1-kHz -1 dBFS signal

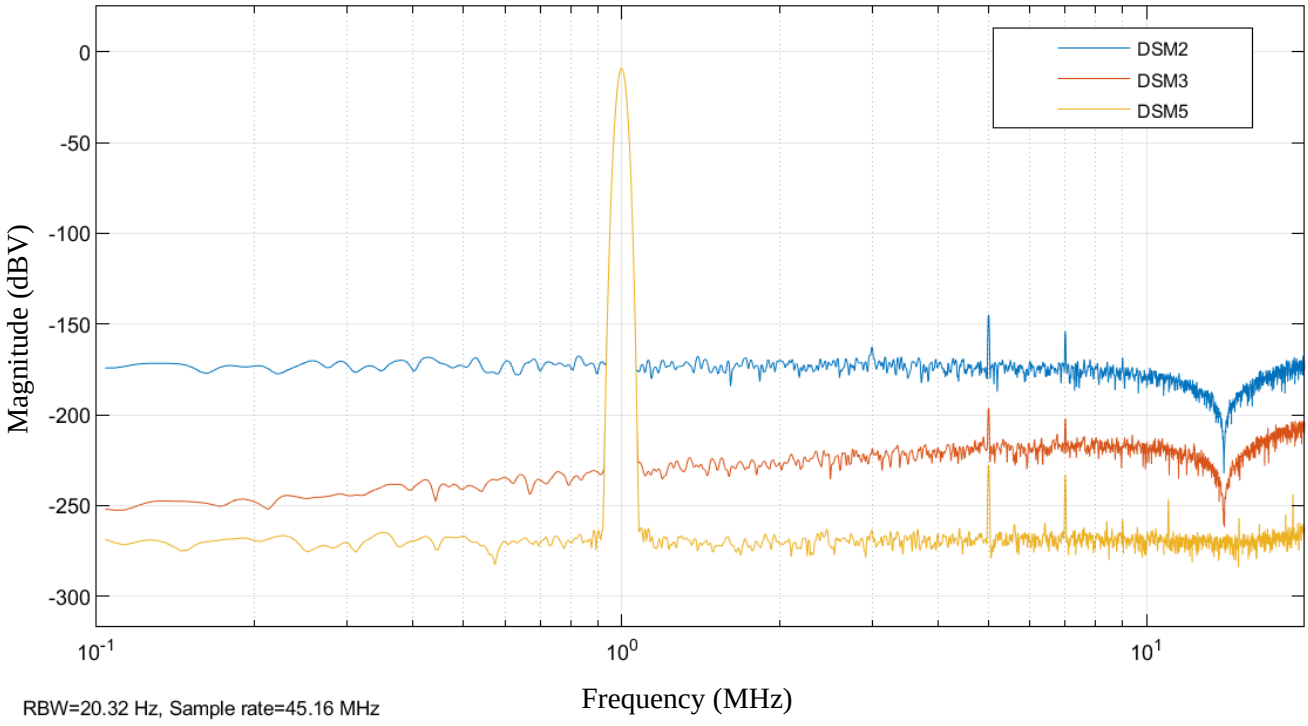


Figure 39 – In-band spectrum of CRFB ΔΣ modulators with 1-kHz -1 dBFS signal

The in-band modulator performance can be more closely examined in Figure 38 for the CIFB loop and Figure 39 for the CRFB loop, respectively. Table 6 summarizes the CIFB loop performance in the 20 Hz – 20 kHz audio band, while Table 7 compares it to the CRFB loop performance.

Table 6 - CIFB ΔΣ modulators performance in the 20 Hz – 20 kHz band

	CIFB DSM2	**CIFB DSM3**	**CIFB DSM5**
THD	-149.57 dBc	-194.97 dBc	-245.53 dBc
SNR	133.59 dBc	168.58 dBc	225.25 dBc
SINAD	133.49 dBc	168.57 dBc	225.21 dBc
SFDR	150.25 dBc	185.81 dBc	229.66 dBc

Table 7 - CRFB ΔΣ modulators performance in the 20 Hz – 20 kHz band

	CRFB DSM2	**CRFB DSM3**	**CRFB DSM5**
THD	-135.11 dBc	-185.84 dBc	-217.12 dBc
SNR	136.38 dBc	175.94 dBc	227.81 dBc
SINAD	132.69 dBc	175.52 dBc	216.77 dBc
SFDR	135.74 dBc	187.07 dBc	218.23 dBc

While the CRFB modulator increases the SNR (even by more than 7 dB in the case of the third-order modulator), it significantly degrades the THD and SFDR performance. The more aggressive noise shaping probably causes more instability due to insufficiently dithered quantizers [12].

The ENOB can be calculated from the SNR as

$$ENOB = \frac{(SNR_{max} - 1.76)}{6.02} . \quad (2.17)$$

Table 8 compares the modulators ENOB. Except for the third-order modulator, the increase in ENOB by using the CRFB loop NTF is insignificant.

Table 8 - The ΔΣ modulators ENOB

	CIFB ENOB	CRFB ENOB
DSM2	21.90	22.36
DSM3	27.71	28.93
DSM5	37.12	37.55

This thesis 18-bit ENOB target is exceeded by all modulators by at least 2.9 bits and therefore demonstrating that even a second-order modulator can achieve performance better than most common DACs if running at high OSR.

None of the modulators showed sings on any idle tones (limit-cycles) under any of the test conditions. The test conditions also included constant DC input sweeps with long steps and signals and noise with different amplitudes and frequencies.

2.2.4 FPGA Implementation

As mentioned above, the modulator cores are generated from MATLAB as VHDL code. The quantizer, feedback and dither are written around the MATLAB HDL cores. Only one modulator block pair (for left and right channel) is active at one time, other modulators are disabled and should not consume any dynamic power. Table 9 compares the FPGA resource utilization of one ΔΣ modulator instance (one channel).

Table 9 - The ΔΣ modulators FPGA resource utilization for one channel

	DSM2	DSM3	DSM5
LUT6	353	394	617
FF	101	149	246

Figure 40 is a snapshot of the FPGA die floorplan after implementation. The reconstruction filter is highlighted in pink, the RF filter in green and the six ΔΣ modulators are highlighted in blue. The filters are placed mostly around the DSP48 and BRAM columns, as expected. The modulators are implemented in generic SLICEx logic. Table 10 summarizes all the Artix 35T FPGA resources used in the DAC implementation.

Table 10 – FPGA resource utilization project summary

Resource	Utilization	
LUT6	4235/20800	20.36 %
of which LUTRAM	940/9600	9.79 %
FF	4234/41600	10.18 %
BRAM	45/50	90 %
DSP48	70/90	77.78 %

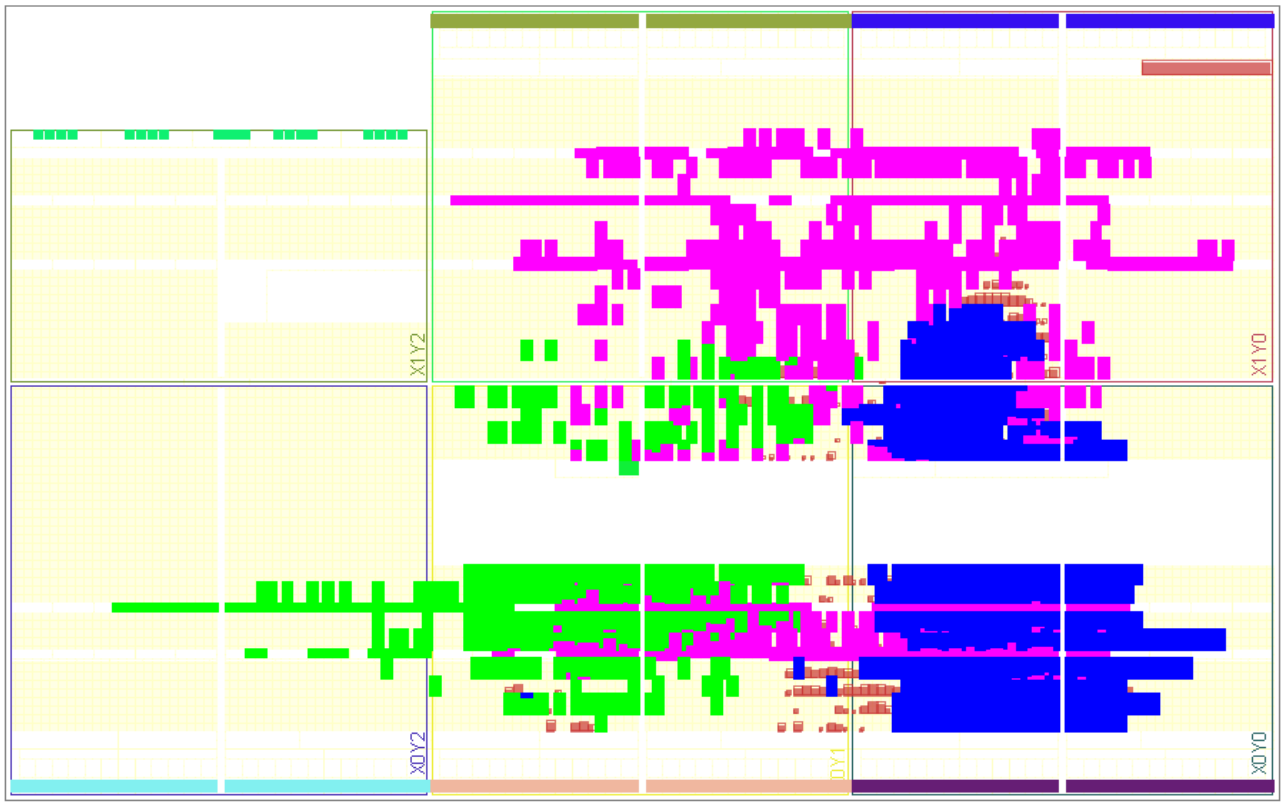


Figure 40 – FPGA die floorplan

3 HARDWARE DESIGN

A hardware platform was created to test the implemented DAC. Digilent ARTY was chosen as the FPGA board. The designed hardware DAC board plugs into the side PMOD connectors on the FPGA board. Figure 41 is a DAC assembly photograph.

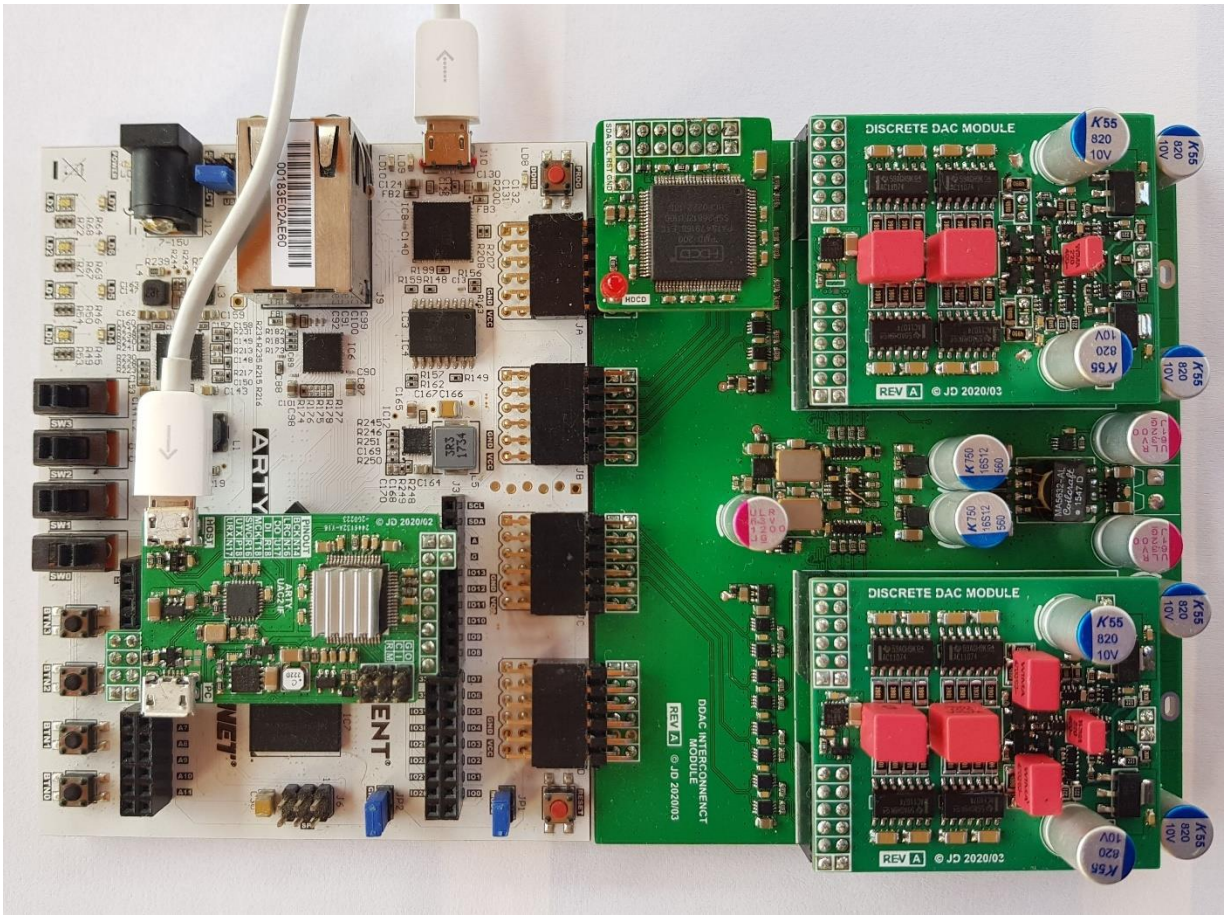


Figure 41- Photograph of the complete DAC board assembly

The USB microcontroller handling the audio data transfer over the USB Audio Class standard is populated on the green board on top of the white FPGA kit. The whole board assembly can be connected by a single USB cable, as there is also an integrated USB hub.

The main DAC board is plugged in over the four PMOD headers on the side of the FPGA board. The two middle PMOD connectors are differential and transfer the four differential signals for the DAC array. The other two PMOD connectors act as control

and power interfaces. The DAC boards itself is a separate PCB assembly plugged into the interconnect mainboard. The mainboard is handling the power supplies and clock distribution. Block diagram of the hardware implementation is shown in Figure 42.

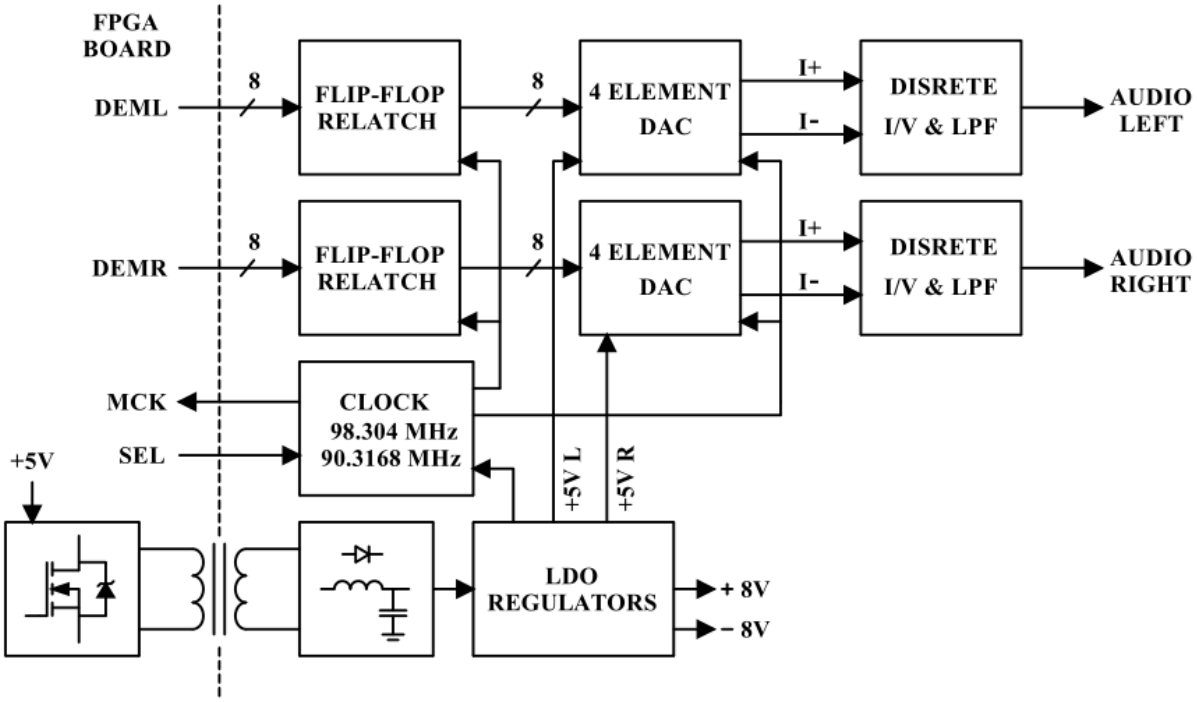


Figure 42 - Block diagram of the DAC hardware

3.1 PDM DAC

The output DAC is one of the most critical parts of this ΔΣ converter. Its purpose is to convert the PDM signal from the modulator into an analog signal. A choice had to be made between single-bit and multi-bit output quantizer (and, therefore, DAC).

Single-bit quantizers have near-perfect linearity, and they are very simple. However, they cannot be adequately dithered [12], making the cyclic stability issues of ΔΣ modulators worse and, therefore, magnifying the idle tone issues. One-bit DACs are also very sensitive to clock jitter. Those are some reasons why most of today's ΔΣ DACs are using multi-bit outputs [1].

This thesis chose a 5 step quantizer and output DAC array. A unit-element DAC was implemented in hardware using four identical elements. The resulting DAC has four levels and a zero (2.2 bits) and can operate at up to 100 MHz. Unit-element DAC linearity problems are solvable by the use of simple techniques such as *Dynamic Element Matching (DEM).* DEM converts the element mismatch error into noise by randomly multiplexing the input data between the DAC elements. The level of noise created by the DEM technique should be minimal as there should not be much element mismatch with 0.1 % thin-film resistors. DEM was chosen over DAW or similar mismatch-shaping methods that have their problems with stability and idle tones. Jitter sensitivity is also halving with each added DAC bit, while SNR is doubling. [1]

Pulse Density Modulation or PDM is a natural quantizer output when it operates within the ΔΣ modulator loop. Multiple methods can be used to convert PDM into an analog signal, most popular being *switched-cap* or *switched-currents* DACs. Both are outside the scope of this thesis as they have to be realized with ASICs.

Creating current or voltage PDM pulses and then integrating them with a *low-pass filter* is a more straightforward solution. Shown in Figure 43 is an example of this method.

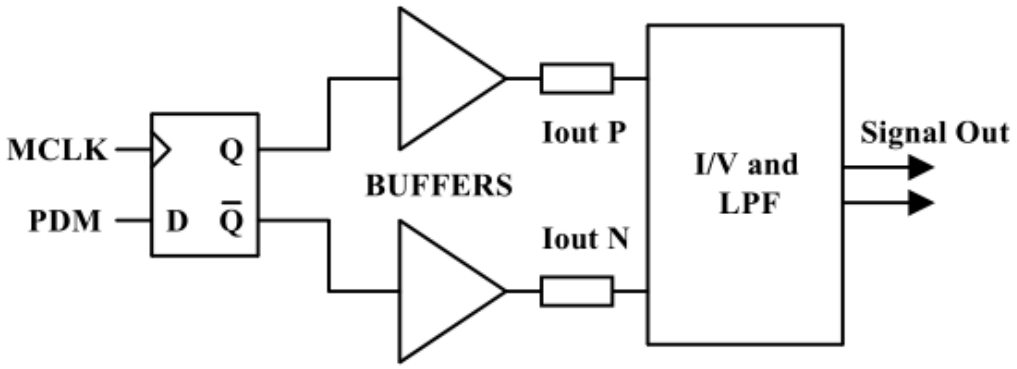


Figure 43 – A simple one-element PDM DAC.

Figure 43 is a fundamental concept of an output PDM DAC. Pulse Density Modulated signal is converted into differential current pulses with logic buffers. These pulses are then integrated with a low-pass filter and converted to voltage. However, this method presumes that the PDM pulses are created ideal and symmetric, as seen in Figure 44A. That, unfortunately, is not the case, and real-world circuits make the pulses non-symmetric as they have limited edge slew rates, parasitics and device manufacturing limitations. These problems can be solved with the *Return to Zero* circuit.

3.1.1 Return to zero

With the simple DAC mentioned above, the PDM pulse areas are integrated by the low-pass filter. This pulse energy must be the same for each pulse; otherwise, distortion will occur. When the modulator outputs two or more consequent pulses, it outputs logic '1' for all consequent clock edges, and thus one long pulse is created. Such behaviour is more similar to low-resolution PWM than PDM. The energy of this multi-pulse is equal to the multiple of single pulses only in ideal conditions.

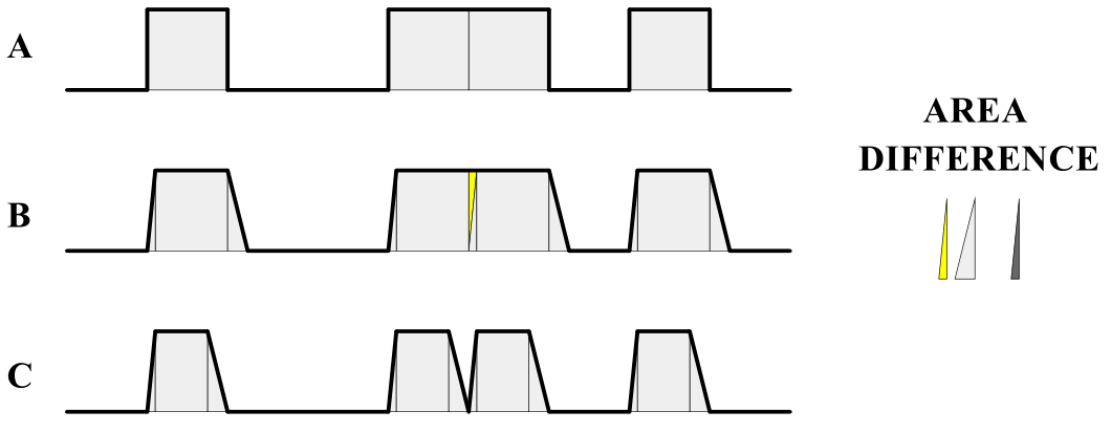


Figure 44 - (A) Modulator PDM output, (B) Edge limitations (C) Return to zero illustration

However, as mentioned in the previous paragraph, the pulses are burdened with edge slew rate limitations and non-symmetricity. Distortion would occur if this “multi-pulse” were to be converted to analog directly due to the areas of the pulse not being precisely equal to a sum of multiple single pulses

This situation is illustrated in Figure 44B. Pulses are illustrated with non-symmetric slew rate limited edges for simplicity. It should be observed that the area of the double pulse is lower compared to two single pulses. This situation can be resolved with the Return to Zero (RtZ) circuit. RtZ circuit will ensure that pulses are always created as single pulses by shortening their duration by one Masterclock (MCLK) cycle, as seen in Figure 44C. One master clock cycle is enough to ensure that the pulse returns to zero before another pulse is created, and therefore no intersymbol interference distortion is created. An example RtZ circuit is shown in Figure 46. This circuit also makes sure that no *common-mode modulation* is created.

3.1.2 Common-mode Modulation Mitigation

Directly applying return to zero to the PDM stream will result in the creation of common-mode modulation. Figure 45 shows signal timing with simple PDM DAC. Similar to Figure 44, but with a return to zero applied. If the output signal is measured differentially, the output value is correct. However, when looking at the common-mode component of each output phase, the signal is not a constant bias, but rather a bias modulation caused by return to zero on the differential phases. In ideal conditions, only the differential signal creates the effective output, and common-mode modulation would not have been a problem. Nevertheless, all circuits have finite common-mode rejection (60 – 70 dB for a well-matched LPF), and common-mode modulation may cause distortion, intermodulation and increased signal correlated output noise.

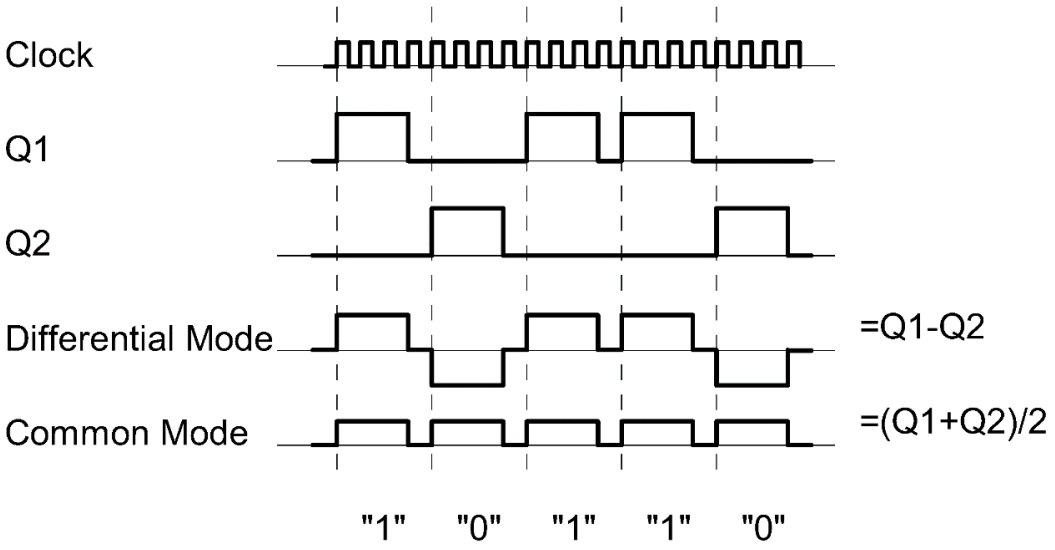


Figure 45 - Common-mode modulation caused by the RtZ circuit

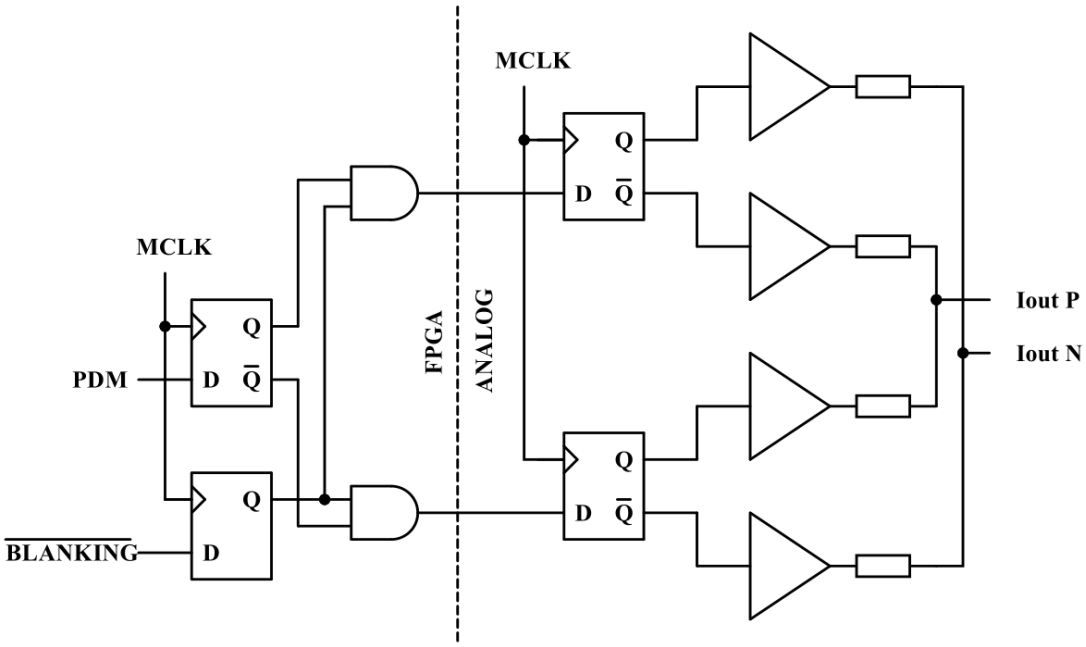


Figure 46 – Circuit to mitigate the common-mode modulation for one PDM bit

The common-mode modulation can be mitigated by adding another opposite phase output element and using a return to zero circuit in Figure 46. The circuit comprises of two D flip-flops, each creating differential phases in its common-mode phase.

The timing diagram in Figure 47 illustrates this case. When the two common-mode phases Q_P and Q_N are summed, the differential output is the same as in Figure 45, only with twice the amount of current. However, the common-mode component of each output phase subtracts. Thus, the result is constant bias current (or voltage) that is not putting stress on common-mode rejection of the LPF filter.

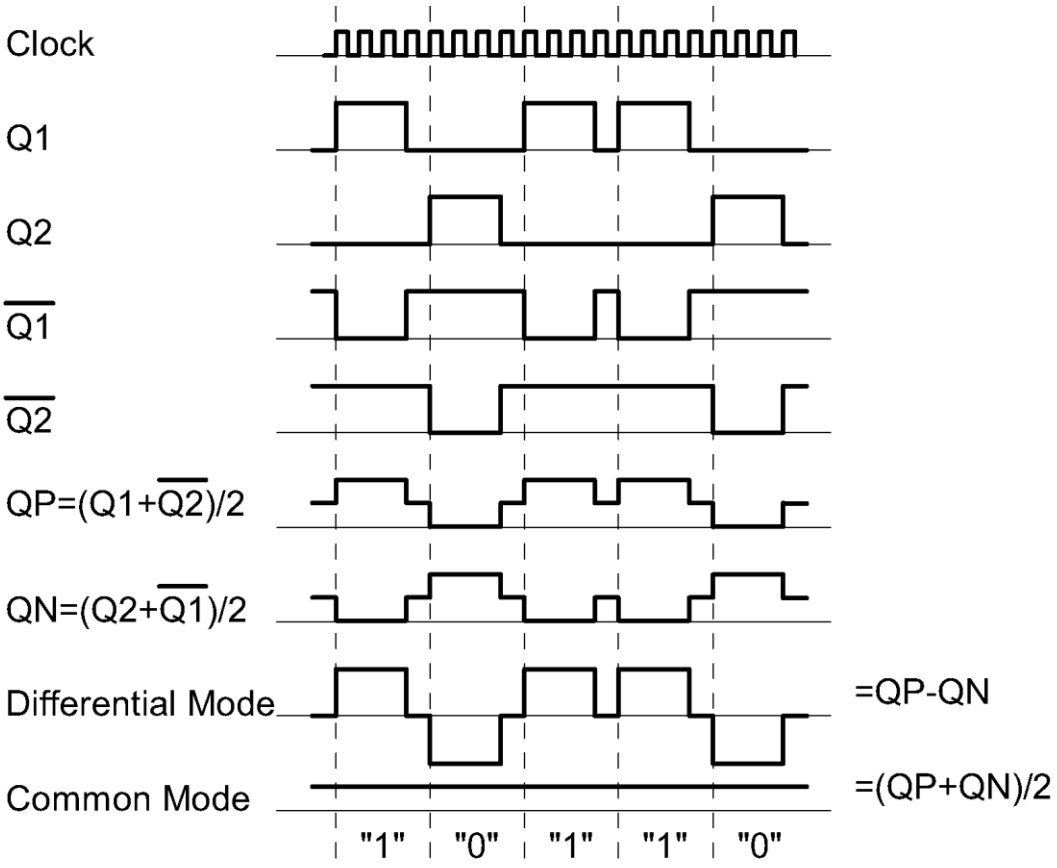


Figure 47 – Mitigated common-mode modulation

The maximum amplitude of the output signal is lowered when applying RtZ, as the pulse area is lower. The ratio of active master clock periods to active periods + blanking periods determines the output amplitude. The ratio of one active edge to two master clock cycles lowers the maximum amplitude by half for a worst-case scenario with OSR of 1024 and MCLK = 98.304 MHz. The amplitude is not lowered as significantly for lower OSR.

3.1.3 Hardware Realization

The DAC itself is realized by a D-type flip-flop TTL logic integrated circuits. Texas Instruments 74AC11074 were chosen for their very high speed and high output current capability, which allows them to drive the resistors directly. The DAC array schematic is drawn in Figure 48.

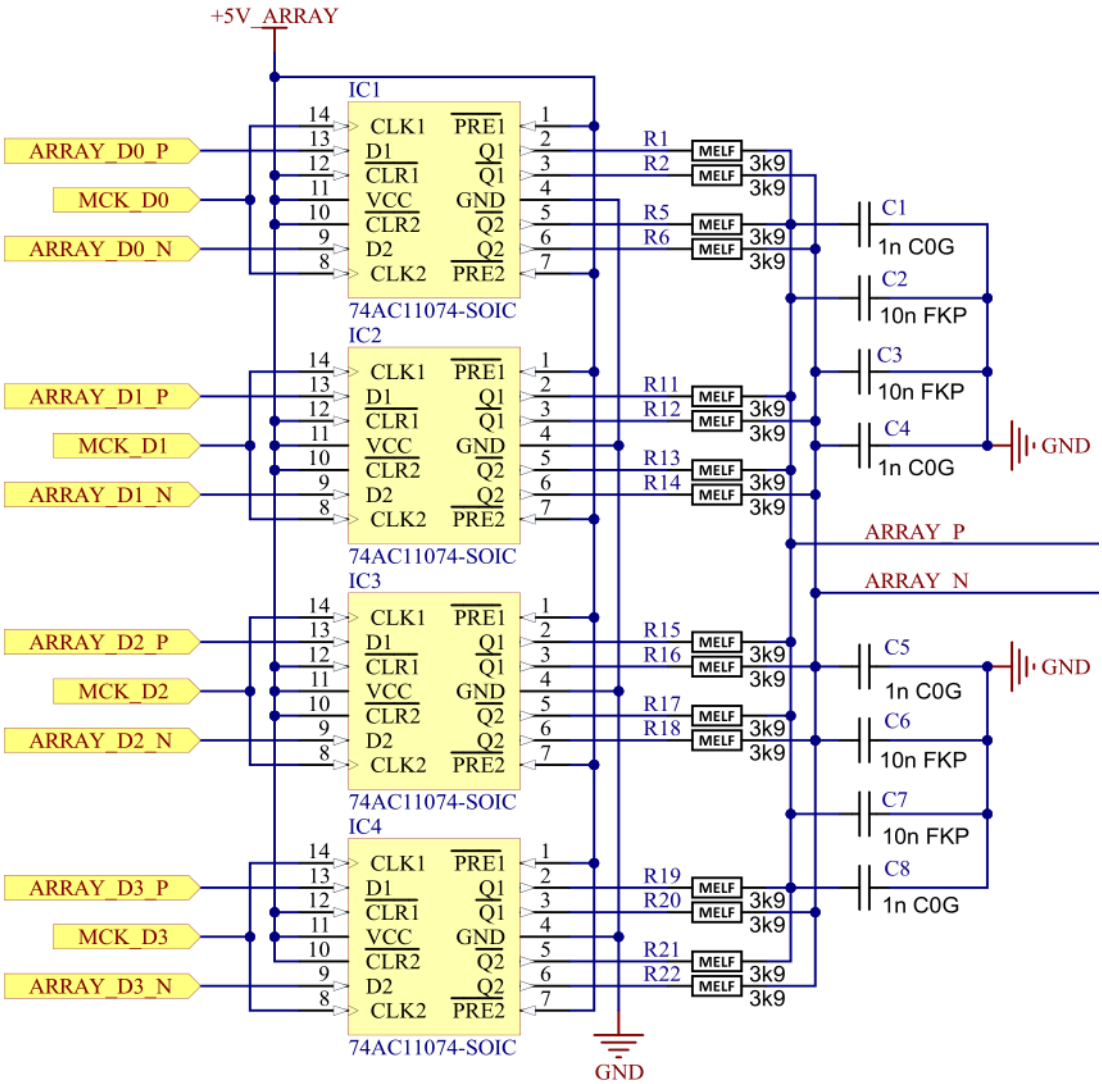


Figure 48 - The 4-element PDM DAC array

The PDM data streams are re-latched by D flip-flops at the main interconnect board before they are routed to the DAC array. The re-latch flip-flops isolate the RF noise and interference from the FPGA. Thus the EMI does not modulate the DAC flip-flops silicon substrate. Each 74AC11074 acts as one element for one PDM data stream. Four elements are needed to build a DAC with five quantizer levels. Precision thin film resistors are placed near the DAC flip-flops. The resistors act as voltage to current converters, and their currents are summed. The low pass filter first stage capacitors are placed near these summing nodes, allowing them to filter the highest frequencies created by the PDM switching.

The resistors are 0.1% thin-film to achieve sufficient element matching. The resistor mismatch would create distortion, but the mismatch is modulated into noise with DEM applied. Lower noise-floor can be achieved with better-matched resistors. Capacitors are a combination of ceramic C0G types and then polypropylene foil dielectrics . Each dielectric is effective at different frequencies.

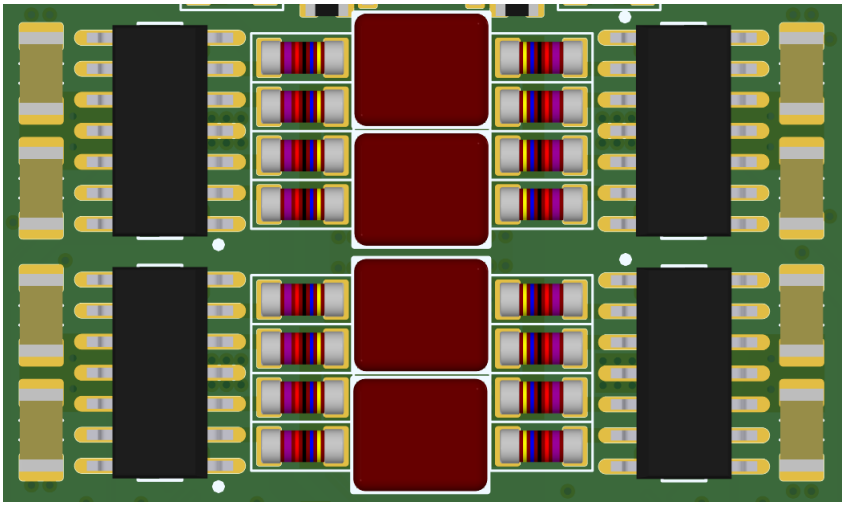


Figure 49 - PDM DAC layout

Figure 49 shows the PDM DAC array PCB layout. Each element has a separate power island with low impedance polymer tantalum capacitors decoupling on the bottom of the PCB. The current outputs sum in the middle node, which has the first stage low pass capacitors in place. Each of the four elements produces 2.56 mA of single-ended current. The whole array therefore outputs about ±10.26 mA of full-scale output current.

3.2 Output Low-pass Filter

The differential current outputs are fed to the output I/V converter and low pass filter. This block is built around a discrete folded-cascode type *operational amplifier (OPA)*. The schematic of the filter is shown in Figure 50.

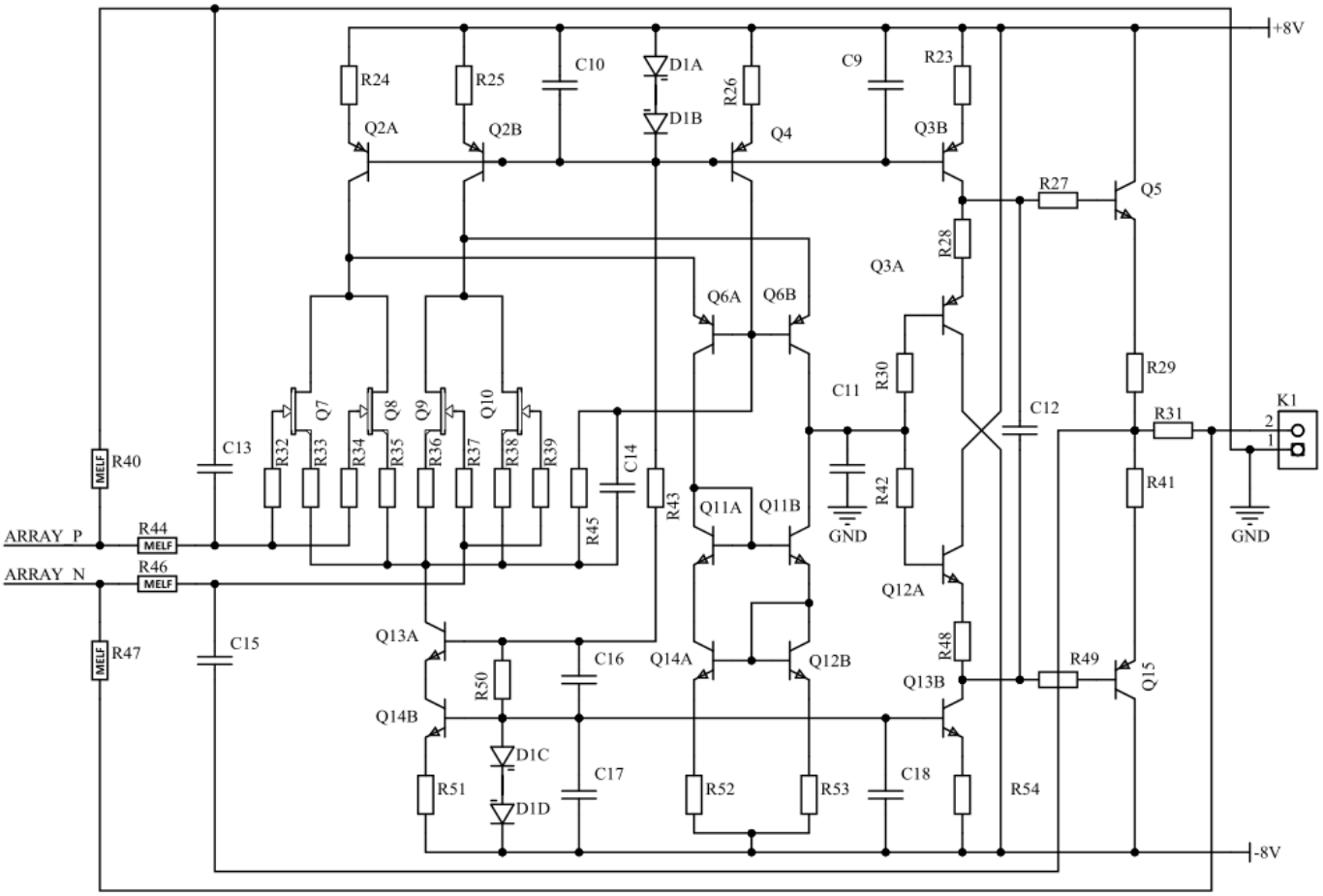


Figure 50 - The discrete output stage OPA with low-pass filter

Table 11 shows some essential parameters of the designed discrete operational amplifier. The OPA is designed to be able to tolerate RF noise from the PDM DAC switching. Therefore a JFET input stage was chosen as JFETs have relatively high EMI immunity. Two low-noise JFET differential pairs are combined in parallel to achieve low noise performance. The output stage can drive headphones directly and remain mostly in class A since it has about 20 mA of bias current. About 8 mA of the bias current is used to drive the feedback to the DAC array.

Table 11 - Output stage OPA parametrics

Parameter	Value
Open-loop gain	100 dB
Open-loop gain at 20 kHz	50 dB
Slew rate	19 V/µs
Phase margin	68°
Systematic DC offset	< 100 µV

The OPA PCB layout can be seen in Figure 51. Mostly dual transistors BJT transistors were used as they provide better matching and space-saving on the board. Low voltage coefficient thin-film resistors were used for critical parts. The input JFET differential amplifier is laid-out in a cross-quad fashion to improve matching and thermal tracking. OPA is on the same PCB as the DAC array for EMI reasons. Four-layer PCB design was chosen.

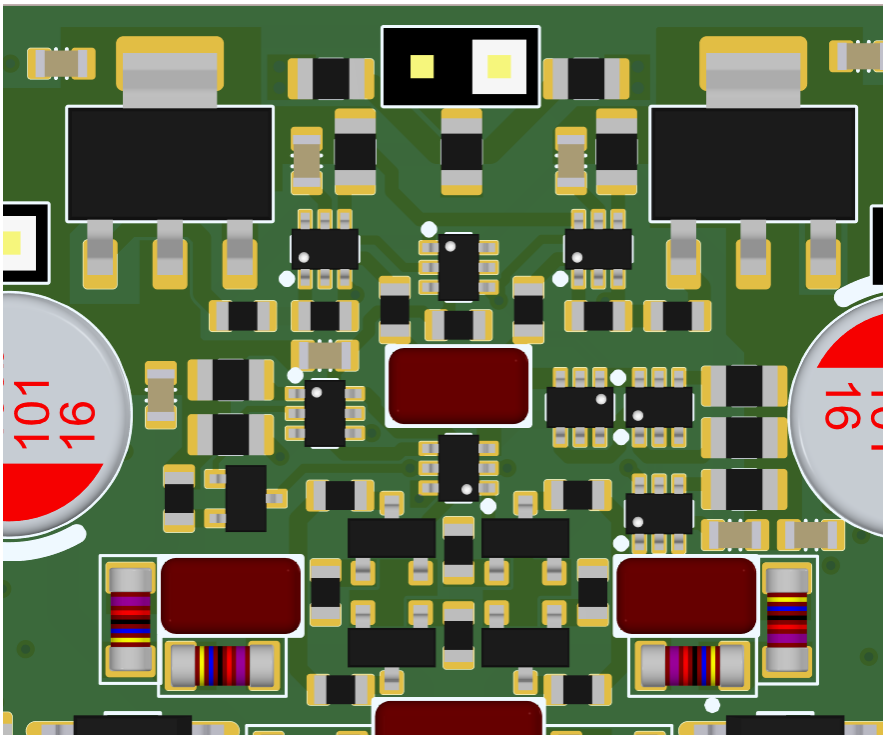


Figure 51 - Discrete OPA output stage PCB layout

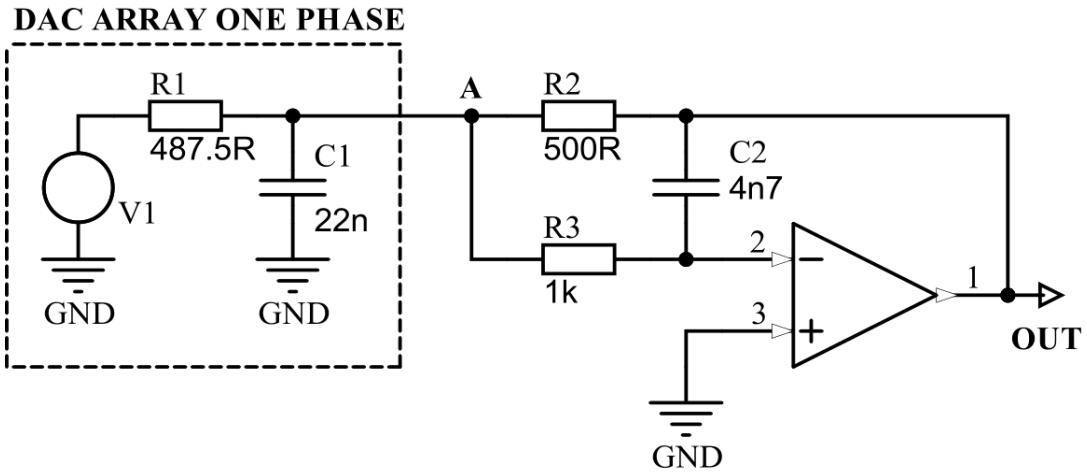


Figure 52 - The analog output stage equivalent circuit

The subsystem of the DAC array and OPA acts as a multiple feedback low-pass filter, with the equivalent circuit in Figure 52. The voltage supply V1 creates the equivalent DAC circuit in combination with resistor R1. The R1 is the DAC impedance. Capacitor C1 is the first stage RF low-pass filtering near the DAC array. Resistors R2, R3 and capacitor C2 complete this circuit into a multiple feedback equivalent second-order low pass filter.

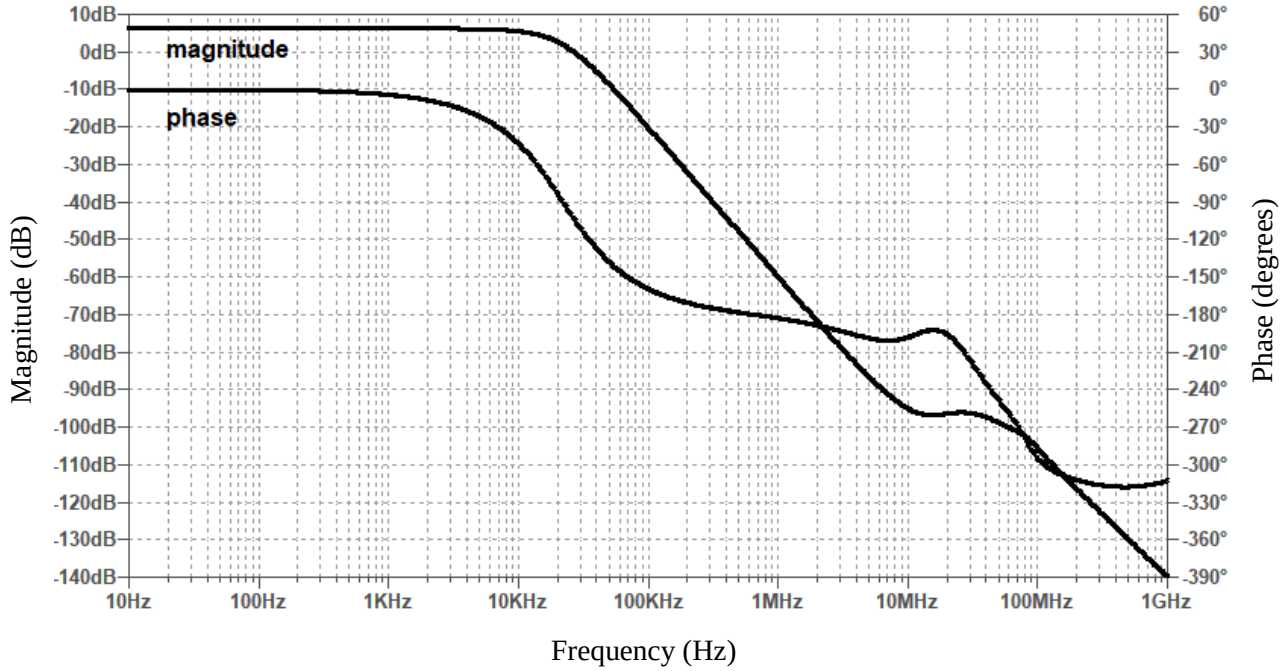


Figure 53 - Magnitude and phase plot of the analog LPF

Figure 53 is a magnitude and phase plot of the analog low-pass filter. The filter has a second-order magnitude response. Bessel approximation was chosen for its constant group delay characteristic in the passband. The filter has a gain of 6 dB, and its -3 dB transition is at about 20 kHz.

The best DAC array operation is ensured if it is driving a zero impedance load, i.e. works in current output mode. Current mode is only possible for the negative array output with the simple differential filter in Figure 47, since the OPA only has a single-ended output to drive the feedback. The positive phase of the DAC array is driving an equivalent 500 Ω load (at DC). This results in non-ideal conditions and higher second harmonic distortion, as described in chapter 4.

The impedance at point A in Figure 52 is plotted in Figure 54. As can be observed, it also is not an ideal near-zero impedance. This is caused by the MFB filter architecture, specifically the R3 C2 low-pass filter in Figure 52. The capacitor C1 takes over at about 20 kHz, and the impedance starts to lower again. Nonetheless, because the positive phase of the DAC array works in voltage output mode, the preference was given to the MFB filter architecture over a simple transconductance amplifier to achieve second-order analog filtering. Table 12 summarizes the designed analog low-pass filter parameters.

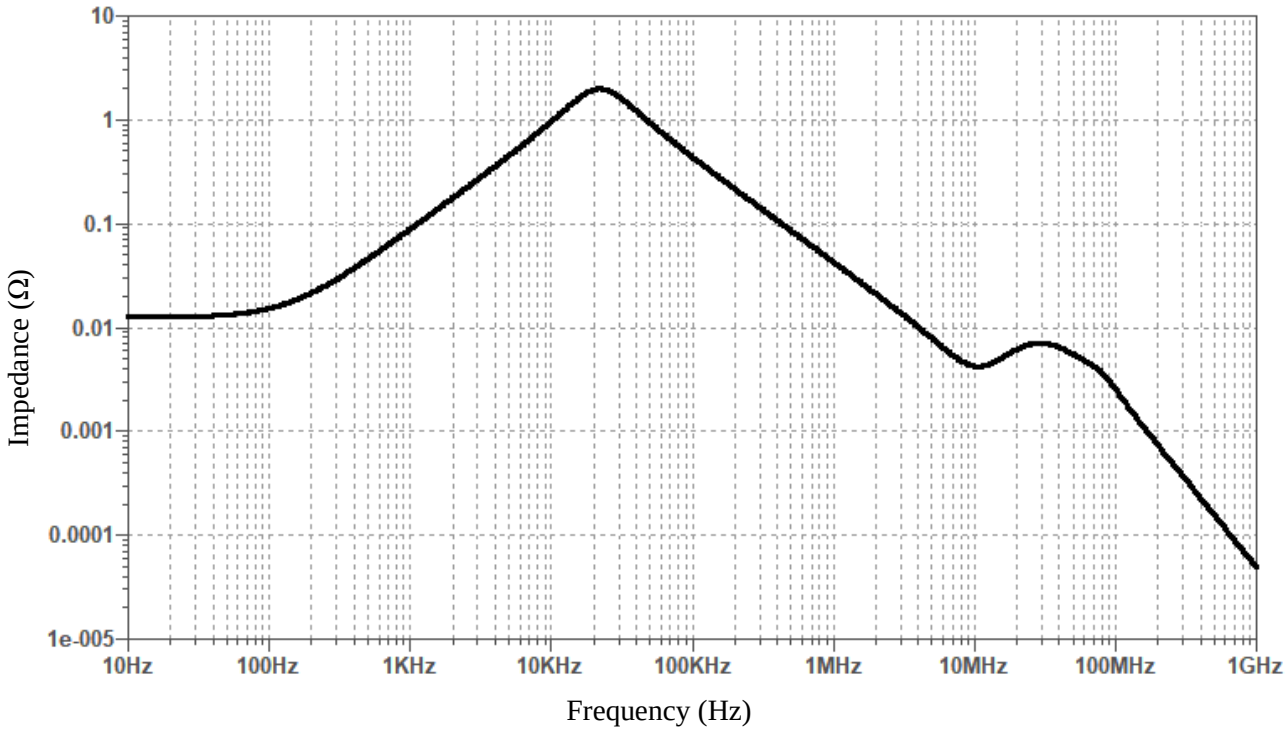


Figure 54 - Impedance plot of the analog LPF at a point A in Figure 48

Table 12 - Analog low-pass filter characteristic

Parameter	Value
Aproximation type	2nd order Bessel
Pass-band gain	6 dB
Pass-band (-3 dB)	DC – 19.8 kHz
Pass-band ripple	None
Atteunation at 10 MHz	> 95 dB

3.3 Power Supplies and Clock

The clock sources are two third-overtone crystal oscillator modules, one with a frequency of 98.304 MHz for 48 kHz input sample rates (48000 FS * 2048 OSR) and the other one with 90.3168 MHz for 44.1 kHz and multiples. Figure 55 shows the clock distribution PCB layout.

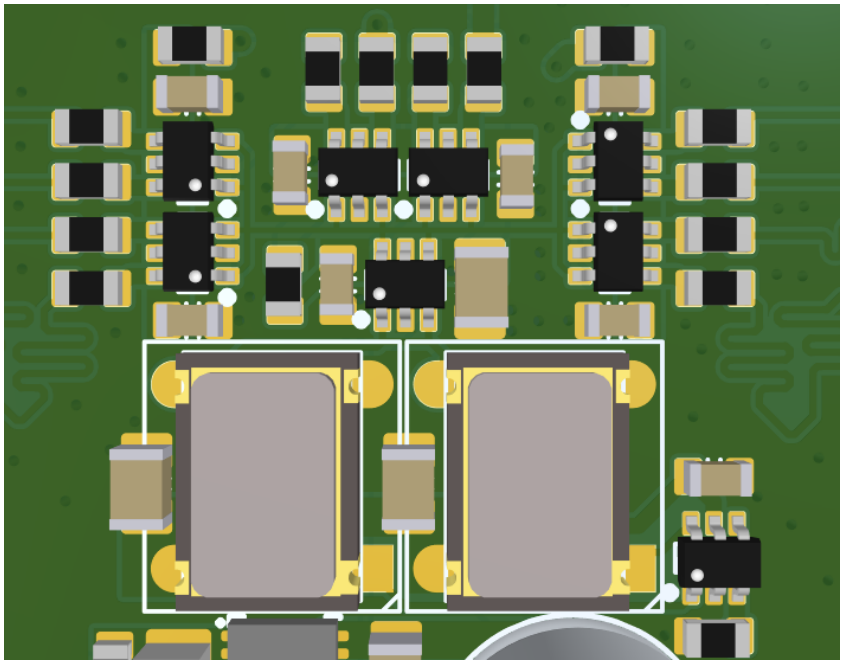


Figure 55 - Clock distribution section PCB layout

The oscillator outputs are buffered and distributed to each part of the system. Each DAC element on each channel has its separate clock buffer to minimize inter-element coupling. TinyLogic NC7WZ04 dual inverters are used for their high speed and relatively low phase-noise.

The distribution blocks have separate power sections for the left and right channels, as any noise on the clock power supply will propagate into the clock outputs as phase-noise. LT3042 was chosen as the LDO regulator to power the clock section since it has low output noise of $0.8\mu V_{RMS}$ (integrated in a band of 10 Hz to 100 kHz). It also has a good *Power Supply Rejection Ratio (PSRR)* of at least 80 dB in the audio band. The clock supply voltage is a 3.65 V to make sure all 5-V supply powered CMOS logic switch reliably due to the CMOS input level threshold.

The same LT3042 LDO IC was used to power each DAC array. The array supply is also a DAC voltage reference, hence a clean, low noise supply is critical. Otherwise, a noise floor of the DAC could be unacceptably high as well as signal modulated, causing noise floor modulation to which the human hearing is very sensitive. The DAC supply schematic is drawn in Figure 56.

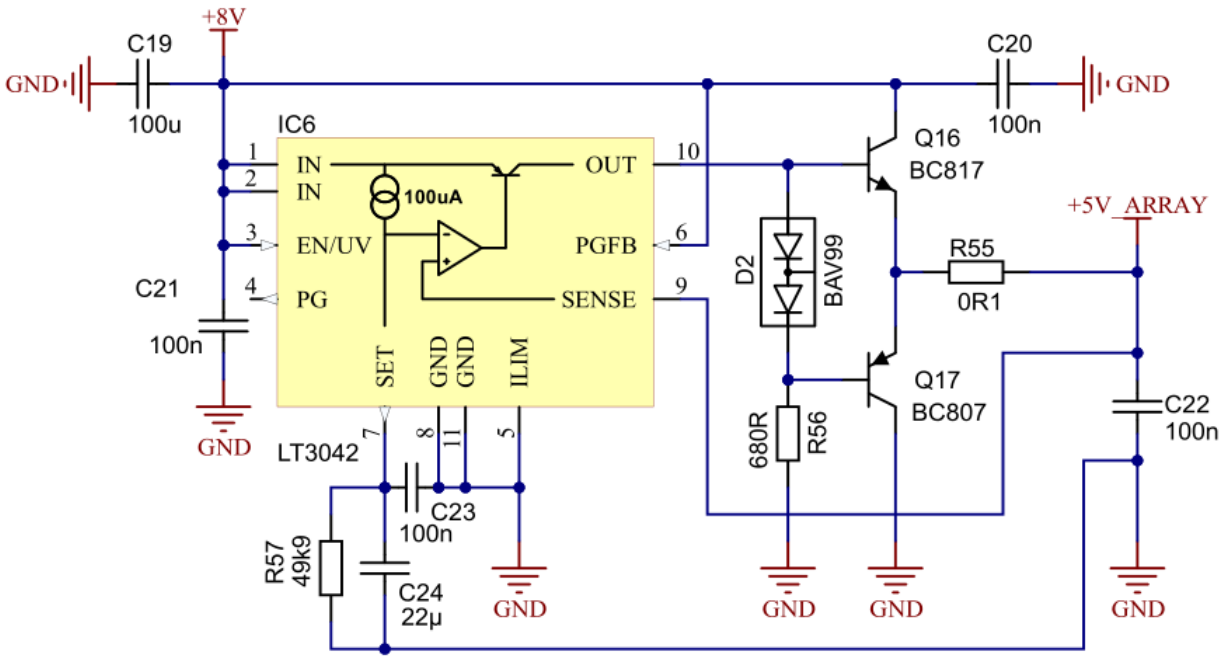


Figure 56 – DAC array power supply

The transistors Q16 and Q17 create class B output stage, allowing the regulator circuit to sink current from the DAC array. Feedback is taken differentially in the middle of the DAC array. The DAC array is also decoupled by 8× 330 µF polymer tantalum low-impedance capacitors in combination with 8× 220 nF X7R dielectric ceramic capacitors.

The operational amplifiers, and the power supplies mentioned, are powered by a symmetric ± 8 V supply. This supply is generated with a feedback-less push-pull type transformer switching power supply, as seen in Figure 57. This switching supply creates galvanically isolated ± 10 V from a 5-V input supply. The LC filtered ± 10-V supply is passed into LT3045 for the positive phase and LT3094 for the negative phase. These LDOs are very similar to already mentioned LT3042, but have a higher current handling capability.

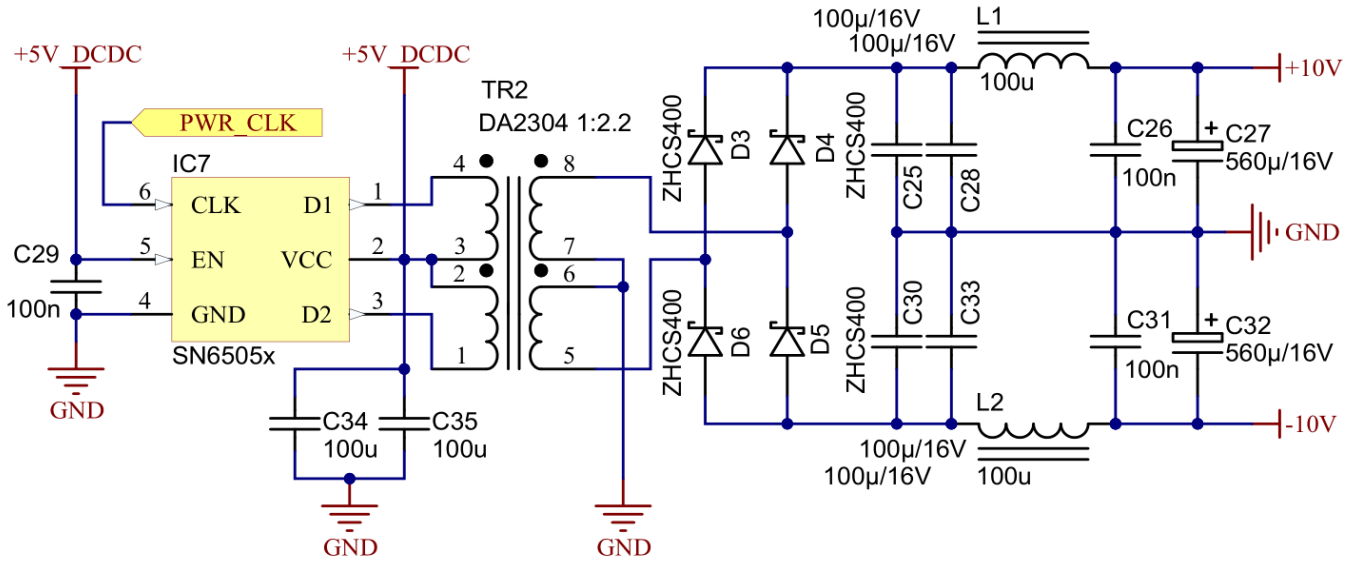


Figure 57 - Switching power supply creating a symmetric galvanically isolated ± 10 V

All switching power supplies are synchronized to the audio clock to limit noise and interference injection into the audio signal. For example, the circuit in Figure 57 is switching at 384 kHz or 352.8 kHz, depending on the input sample rate.

4 MEASUREMENTS

The DAC was characterized using an audio analyzer system described in the bachelor thesis [13]. SpectraPLUS is used as the FFT software.

The analyzer system from [13] uses a -30dB high-Q analog notch filter at the fundamental harmonic. This notch filter drops the input energy into the ADC in the analyzer significantly. Lower input energy into the ADC is consequently nearly eliminating the harmonic distortion of the ADC itself. The higher harmonics amplitudes are kept as created by the *Device Under Test (DUT).* The analyzer can measure below -140 dB THD with 126 dB of SNR.

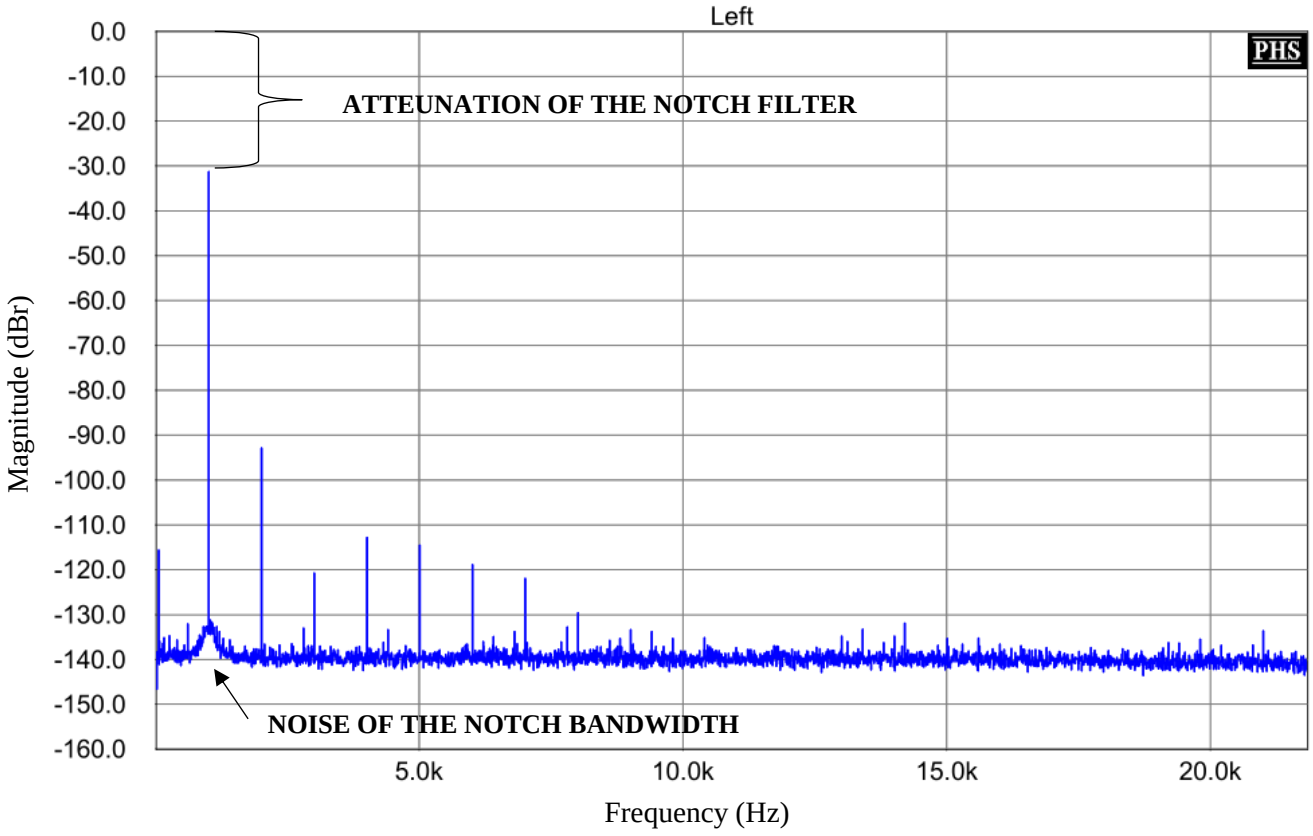


Figure 58 - THD measurement example spectrum

Figure 58 is a THD measurement example made with the analyzer. The notch filter attenuates the fundamental harmonic by about 31.7 dB. The notch filter also adds noise, which can be observed in Figure 58 as well. This noise is limited to the notch filter

bandwidth, and it does not affect any of the higher harmonics. The notch attenuation must be taken into account when calculating the THD.

For all measurements, a 131072 block size FFT with a 48 kHz sample rate is used. The full-scale output amplitude of the DAC is 2 V_{RMS}, corresponding to 0 dBr in the FFT magnitude graphs. Both channels measured identical. Only one channel is shown for clarity. Only CIFB modulators were characterized, as CRFB architecture did not bring any measurable improvement.

4.1 Signal to Noise Ratio

Signal to Noise Ratio is a measure that compares the level of the desired signal to a background noise. Integrating the DAC output noise floor yields an SNR measurement as the 0 dBr point is calibrated to correspond to the DAC full-scale output.

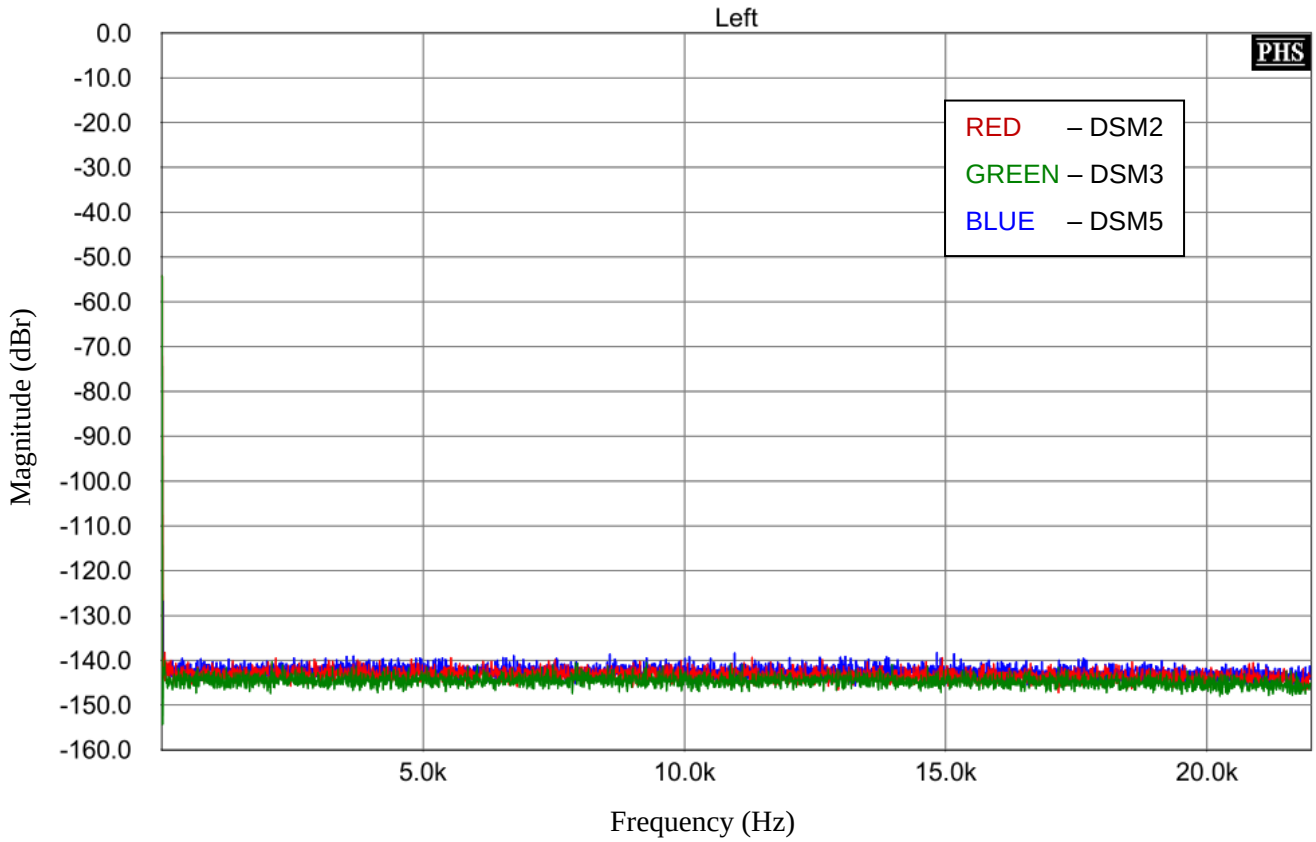


Figure 59 - DAC noise floor with DEM enabled

The noise floors in Figure 59 are a measurement of all three modulators with no input signal, while the dynamic element matching is enabled. Figure 60 shows the same measurement, but with dynamic element matching disabled.

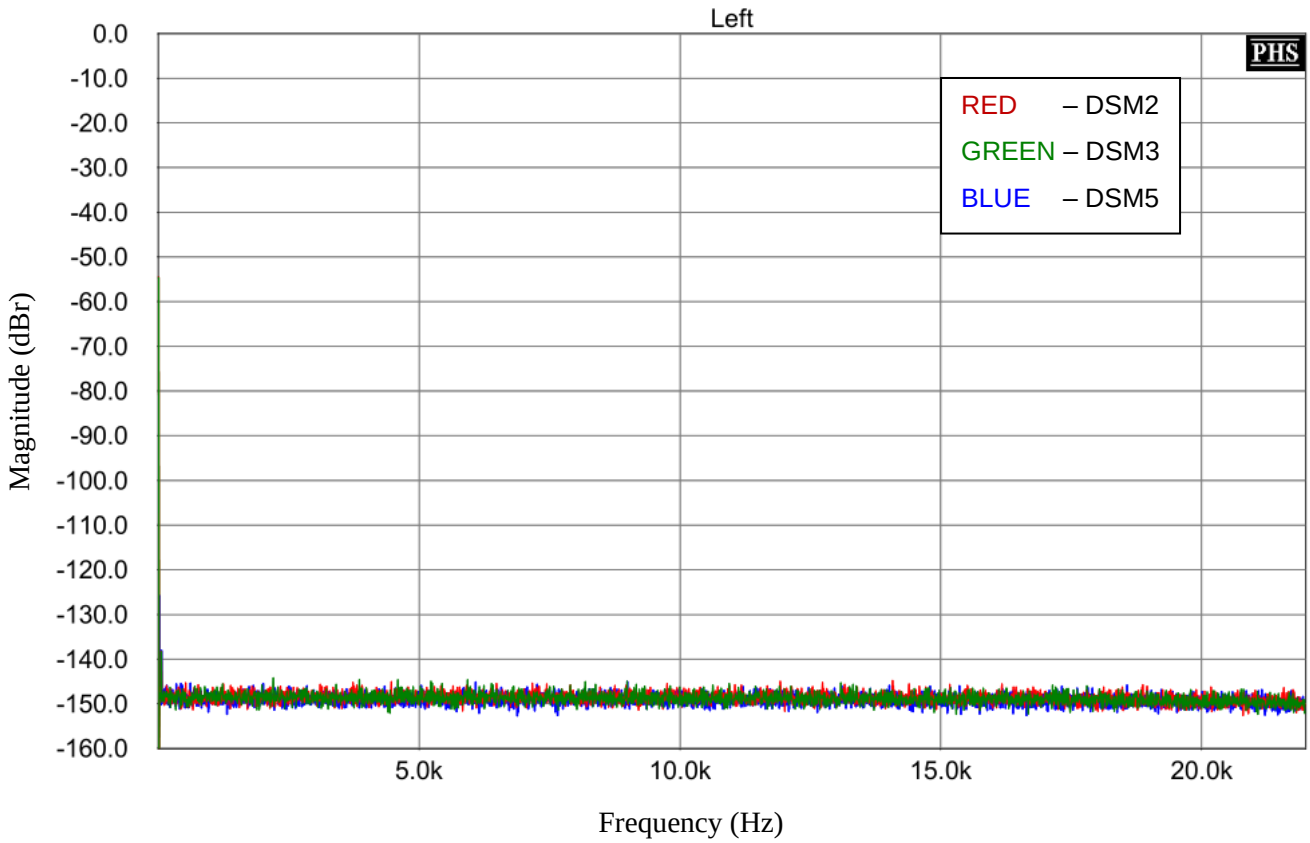


Figure 60 - DAC noise floor with DEM disabled

The noise added by the dynamic element matching is 4.9 dB for all modulators, shown in Table 13.

Table 13 - ΔΣ DAC signal to noise ratio integrated in a 0.37 Hz to 24 kHz band

	DSM2	DSM3	DSM5
SNR w/ DEM	119.8 dBr	119.8 dBr	119.8 dBr
SNR w/o DEM	124.7 dBr	124.7 dBr	124.7 dBr

The noise added by DEM directly correlates with the resistor mismatch in the DAC array. Better than 0.1 % resistors have to be used or more DAC elements have to be implemented in order to achieve better noise performance with DEM enabled.

4.2 Dynamic range

Dynamic range is a ratio between the largest measurable signal and the smallest measurable signal. The dynamic range test is usually based on a THD+N measurement made on a -60 dBFS signal and adding 60 dB to the result. The fundamental harmonic is removed out of the band, and the energy in the band is integrated (form 0.37 Hz to 24 kHz in this case). This integrated spectrum is then compared to the full scale fundamental harmonic, resulting in a THD+N ratio.

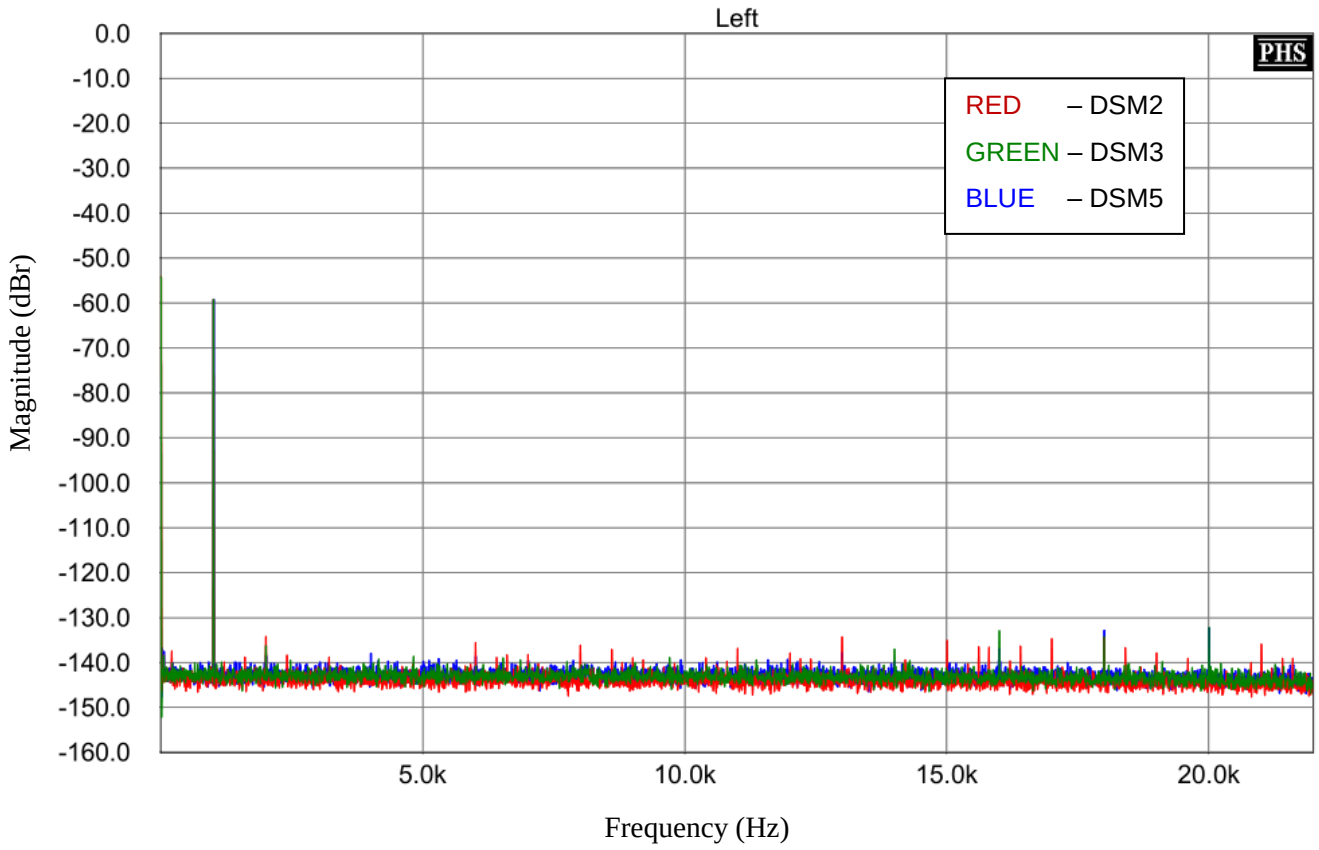


Figure 61 - DAC output spectrum with -60 dBFS input and DEM enabled

Figure 61 is a dynamic range measurement on the DAC with DEM enabled. Low amplitude harmonically related components can be observed for the second and third-order ΔΣ modulators. These harmonics are probably an effect of the noise shaping loop running out of loop gain.

Figure 62 shows the same measurement, but with dynamic element matching disabled.

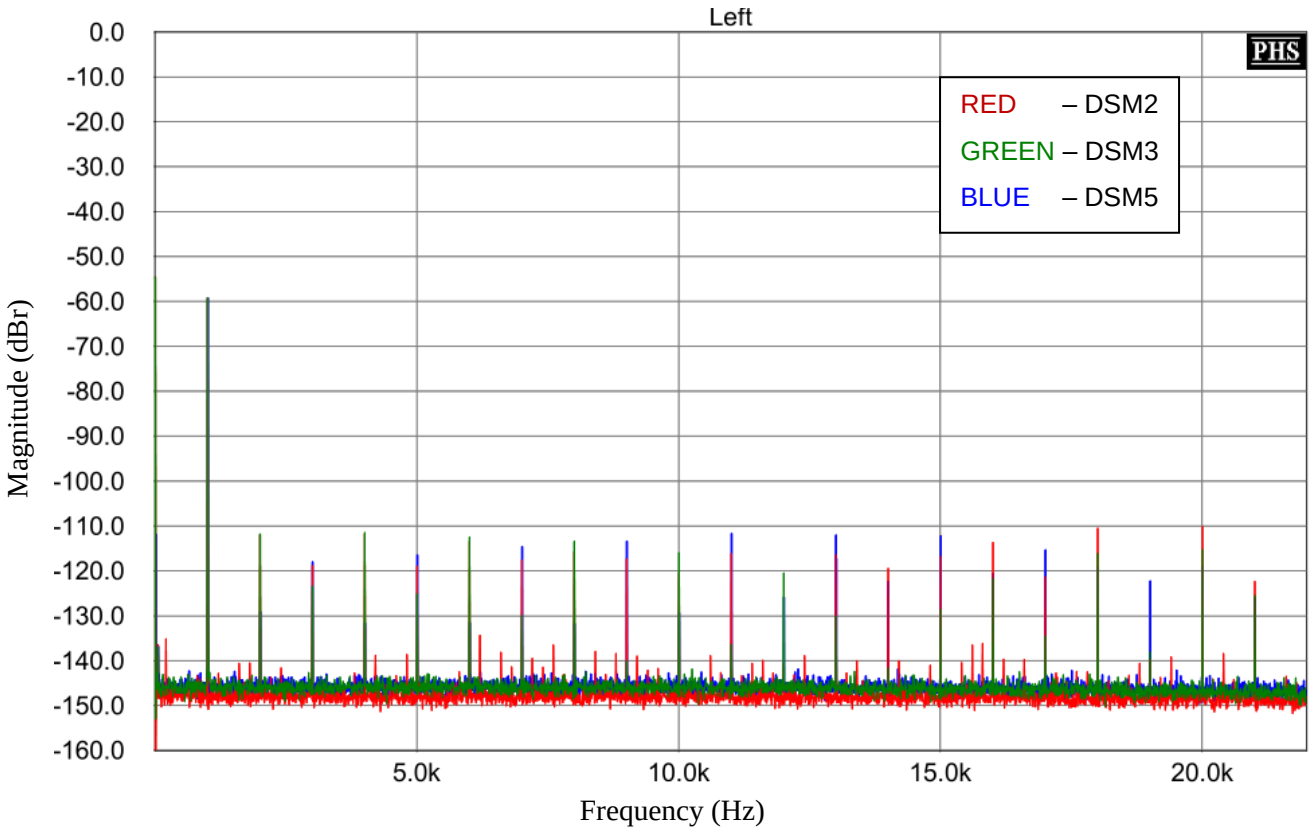


Figure 62 - DAC output spectrum with -60 dBFS input and DEM disabled

The benefit of linearizing the DAC array with dynamic element matching is obvious. The noise floor is lower, but the harmonic distortion caused by the mismatch in the DAC array is very significant. The harmonic distortion caused by the non-linear DAC array far outweighs the detriment of 4.9 dB noise floor increase. This effect is further emphasized by the dynamic range results summarized in Table 14.

Table 14 - ΔΣ DAC dynamic range measured in a 0.37 Hz to 24 kHz band

	DSM2	DSM3	DSM5
DR w/ DEM	119.4 dBr	119.7 dBr	119.8 dBr
DR w/o DEM	104.4 dBr	104.7 dBr	104.6 dBr

Disabling the dynamic element matching reduced the dynamic range by about 15.2 dB in the worst case for the fifth-order ΔΣ modulator. The dynamic range is the same as SNR for the fifth-order modulator. This interesting observation is indicating no noise-floor modulation with a signal. For the second-order modulator, there is an SNR to DR difference of 0.4 dB. However, this is probably caused by the higher-order harmonic components seen in the spectrum, not an actual noise floor modulation. The same is true for the 0.1 dB difference of the third-order modulator.

4.3 Harmonic distortion

Total Harmonic distortion is a ratio between the sum of energies of all higher harmonic components to the fundamental harmonic. The DUT is driven with a full-scale harmonic signal during then measurement. Harmonic distortion directly relates to the system linearity. The THD measurement method used is described in detail in [13]. The short description is at the beginning of this chapter. Each ΔΣ modulator is measured separately for clarity. The DEM, as well as the notch filter, are enabled for all cases.

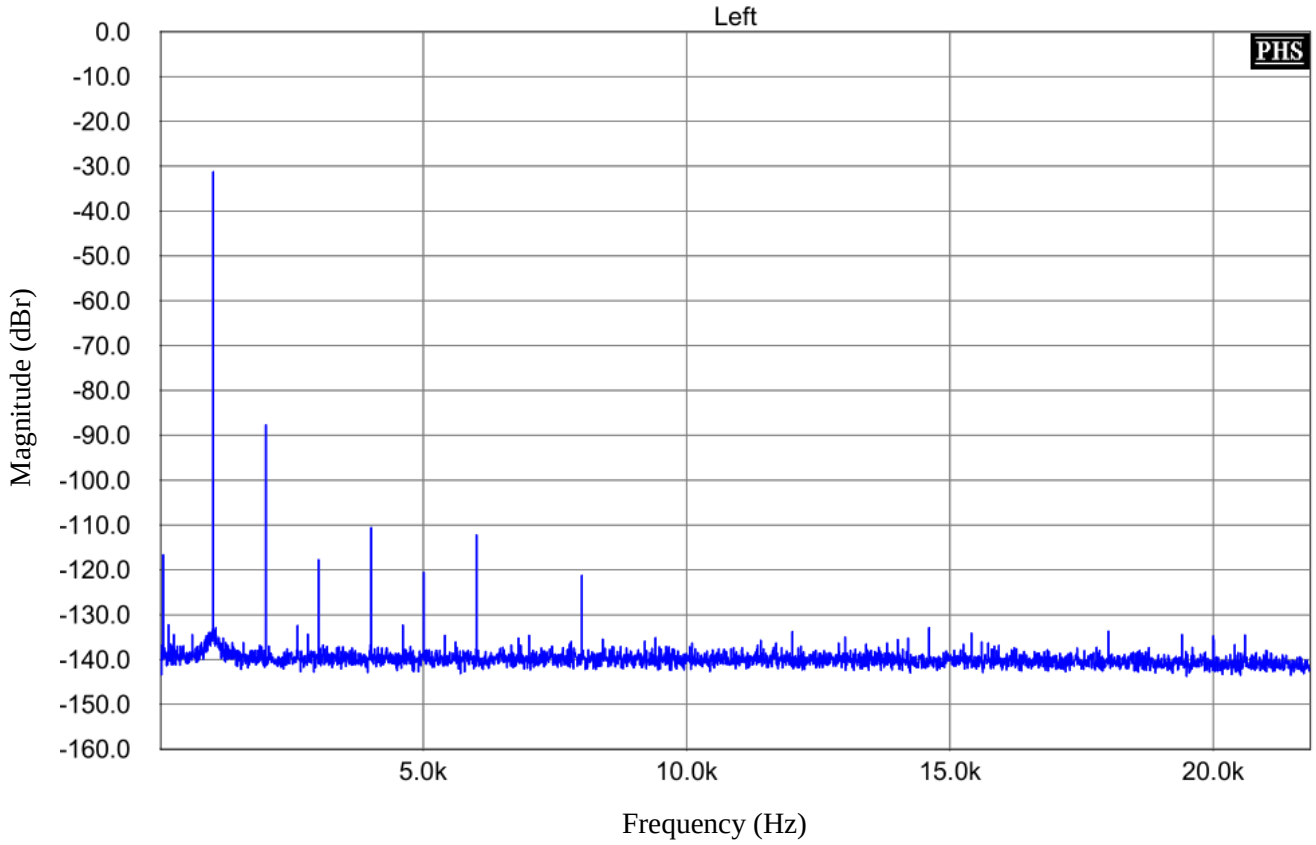


Figure 63 - ΔΣ DAC spectrum for second-order modulator with 0 dBFS 1kHz input

Figure 63 is a DAC spectrum with the second-order ΔΣ modulator enabled and a full-scale 1-kHz sine input. The 1-kHz harmonic signal is attenuated by about 31.7 dB by the notch filter. The rest of the spectrum is untouched. Similarly, Figure 64 shows a spectrum with the third-order modulator enabled.

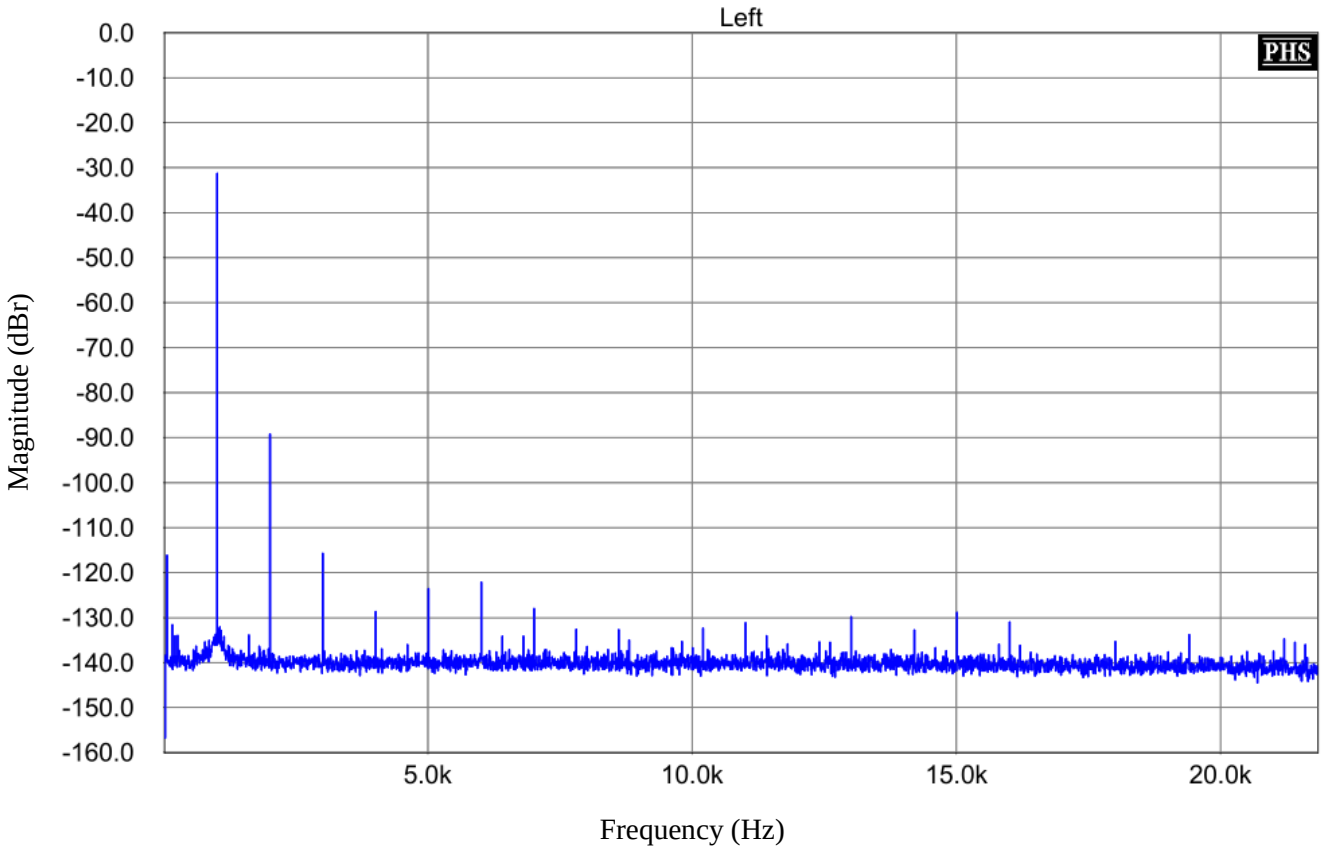


Figure 64 - ΔΣ DAC spectrum for third-order modulator with 0 dBFS 1kHz input

Table 15 summarizes the measured total harmonic distortion for all modulators. Figure 64 is a spectrum for the fifth-order modulator.

Table 15 - ΔΣ DAC total harmonic distortion for 12 harmonic components

	DSM2	DSM3	DSM5
THD (dB)	-88.6 dB	-89.9 dB	-90.2 dB
THD (%)	0.0037%	0.0032%	0.0031%

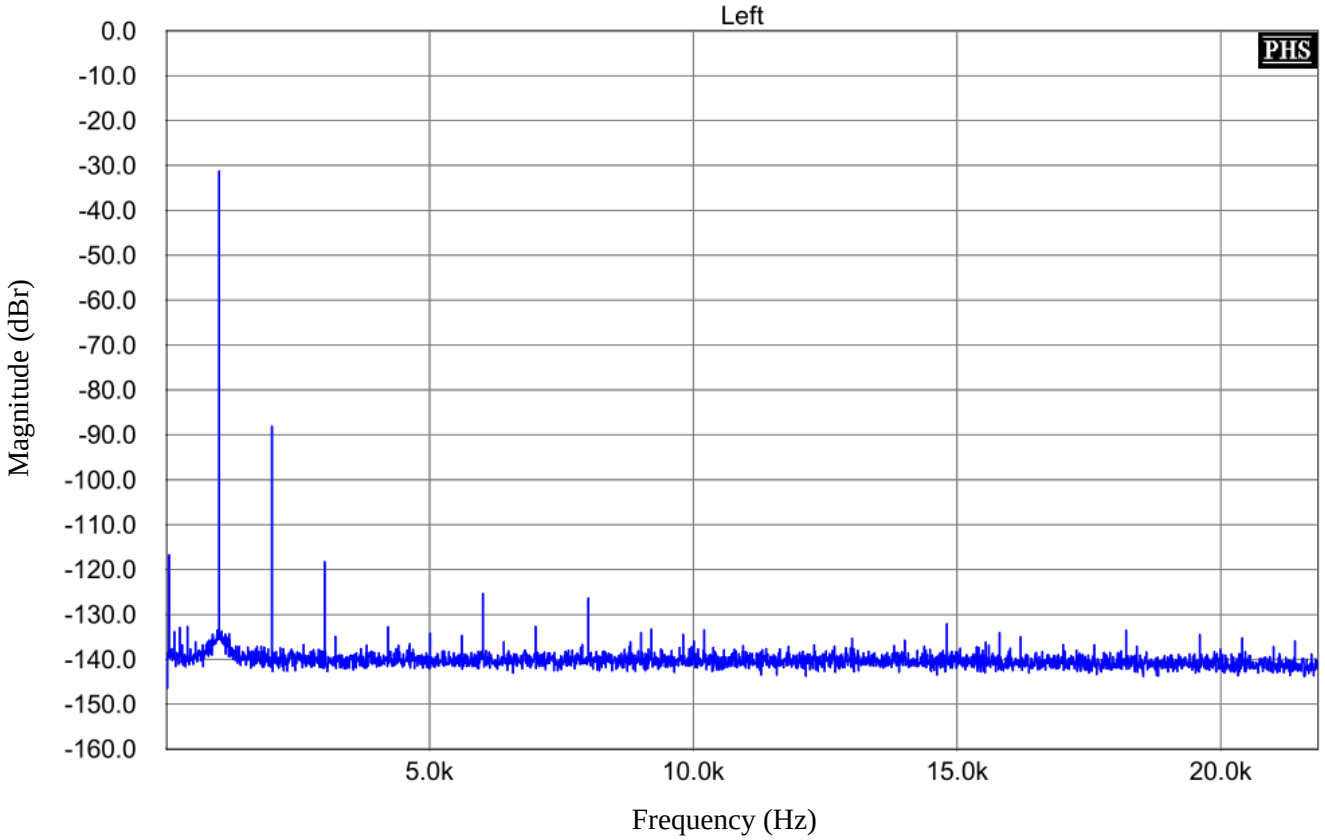


Figure 65 - ΔΣ DAC spectrum for fifth-order modulator with 0 dBFS 1kHz input

Table 15 shows nearly identical THD measurements for the different modulators, pointing to the analog section and the DAC as the limiting factor. The high second harmonic component is indicative of an unbalanced system and is caused by the analog output filter. The output filter loads each phase of the DAC array differently. The negative phase of the DAC array is loaded with low impedance and working in current mode, while the positive phase is loaded by 500 Ω operating in voltage mode, as described in chapter 3.

The load impedance unbalance is the leading cause behind the very high second harmonic distortion. Experiments have shown that the second harmonic distortion can be suppressed with a *fully-differential amplifier (FDA)*. FDA's have two feedback paths allowing both sides of the array to work in current mode, and therefore nearly eliminating the even-harmonics.

4.4 Power and summary

Table 16 summarizes the ΔΣ DAC achieved performance and other characteristics. Power consumption metrics are added. The DAC PCB consumes constant 5 W at OSR of 1024 for two channels. The 5 W power consumption includes the clocks, re-latches, DAC arrays, and output stages. A lot of the power is consumed by the output stages, each taking about 600 mW. Output stages add up to about 2 W when accounting for LDO and switching power supplies efficiency. About 320 mW is consumed by the clock circuit, which drives the DAC elements, re-latches and FPGA. The rest is the dynamic power of the DAC arrays themselves and the re-latching flip-flops.

The power consumption does not change much with different modulators, indicating that most of the power is consumed by the high-speed FIR filter convolution. The finalized DAC consumes between 6.2 W to 6.4 W when the USB interface and rest of the FPGA board power consumption is added. This power consumption is perfectly acceptable as the DAC was never designed for battery operation and would be used most likely in a Hi-Fi system.

Table 16 - ΔΣ DAC performance summary

	DSM2	**DSM3**	**DSM5**
THD (dB)	-88.6 dB	-89.9 dB	-90.2 dB
THD (%)	0.0037%	0.0032%	0.0031%
SNR	119.8 dBr	119.8 dBr	119.8 dBr
DR	119.4 dBr	119.7 dBr	119.8 dBr
ENOB	19.61 bits	19.61 bits	19.61 bits
FPGA core current	673 mA at 0.95 V	685 mA at 0.95 V	702 mA at 0.95 V
DAC current	1 A at 5 V	1 A at 5 V	1 A at 5 V
Total power	6.2 W	6.2 W	6.4 W

5 CONCLUSION

This thesis aimed to design high-OSR lower-order ΔΣ modulators and demonstrate that they can be used as an alternative to more complicated high-order ΔΣ modulators. Basic principles behind ΔΣ modulators, such as oversampling and noise-shaping, were explained. Quantizers were briefly discussed as they are a critical part of any ΔΣ converter. The thesis followed by an overview of single loop ΔΣ modulator architectures.

An interpolating reconstruction filter was designed with an emphasis on time-domain performance, which was achieved with a custom windowing function. The reconstruction filter is implemented as a polyphase FIR convolution with 32000 coefficients. The reconstruction filter output is a 16× oversampled input signal. This signal is further interpolated by a factor of 8 by the RF filter, removing spectral images up to 128× of the input rate.

Three ΔΣ modulators were designed. Each has a resonator loop capable of creating a zero in the in-band spectrum and extending the modulators SNR. The resonator blocks could be disabled to allow for a comparison between the CIFB and the CRFB modulator architecture. Second-order, third-order and fifth-order modulators were designed, all running at OSR of 1024. The second-order modulator achieved SNR of more than 133 dBc in simulation, while the fifth-order modulator achieved over 225 dBc of SNR. No noise-floor modulation or idle-tones were observed.

The filters and modulators were implemented into an Artix 7 35T FPGA. Supporting blocks, such as I2S input, audio decoder, dither generators, DEM or RtZ, were also implemented. VHDL is used for the implementation. The modulator loop transfer functions were generated by a MATLAB Simulink HDL Coder. Finished implementation consumed about 20 % of the FPGAs LUTs, 10 % of flip-flops, 80 % of DSP48 blocks and 90% of BlockRAMs.

Digilent ARTY was chosen as the FPGA development board. The DAC hardware was designed to be plugged into the FPGA board. As switched-cap or switched-current DACs require ASIC design, a more straightforward solution with direct four-element PDM conversion was chosen. Firstly, the Return to Zero circuit formats the PDM pulses to keep the pulse energy constant. The processed PDM pulses are then converted to an

analog signal with a 2.2-bit unit-element DAC running at nearly 100 MHz. One DAC element is comprised of dual high-speed, high output current D flip-flop IC. Each of the four elements is driving four precision thin film resistors, resulting in 16 precision resistors acting as voltage to current converters. Resulted currents are differentially summed. A dynamic element matching method is used to convert the resistor mismatch into noise, linearizing the DAC significantly.

These pulses are integrated by a second-order low-pass filter with Bessel approximation. At the core of this filter is a discrete folded cascode operational amplifier with high output current capability, allowing to drive headphones directly. The discrete OPA also has a cross-quad style low-noise JFET input stage as JFETs have higher EMI immunity.

Low noise power supplies were designed. Noisy voltage reference will cause noise floor modulation and increased noise-floor of the DAC. Therefore, a very low noise supplies were designed for the DAC array, since the DAC power supply is also a voltage reference. Clock generator and distribution block with is own low noise power supply was also designed.

An audio analyzer described in [13] was used to characterize the ΔΣ DAC. A SNR of 119.8 dBr was achieved for all modulators with DEM enabled. This SNR is equivalent to 19.6 ENOB, exceeding the assigned 18 ENOB performance. The dynamic ranges measured were 119.4 dBr for the second-order modulator, 119.7 dBr for the third-order modulator and 119.8 dBr for the fifth-order modulator. A quite high THD was measured at about -90 dBr for all modulators. The high THD is caused by a high second harmonic component, created by a non-balanced DAC array load. The positive and negative output phases are loaded differently due to the single-ended low-pass filter architecture, causing this high second harmonic distortion. A fully-differential output filter can achieve about 30 dBr lower total harmonic distortion.

Nonetheless, the measurements have proven that even second-order ΔΣ modulator is sufficient to reproduce audio signals in high fidelity, as long as it runs at high oversampling rates. Third and fifth-order modulators do not bring much benefit, as the performance is limited by the analog stage.

5.1 Follow-up work

Many aspects of this ΔΣ DAC can be further improved. The most significant improvement can be made by redesigning the output low-pass filter to eliminate the second harmonic distortion. A fully differential OPA has to be designed. More DAC elements will improve the SNR and DR by spreading the matching error.

The output quantizer can be forced to work in a PWM mode by injecting a saw-tooth wave into the loop before the quantizer. PWM mode may bring many benefits. Namely, the number of switching transitions in the DAC array would be constant regardless of the input signal and would not be causing any signal related modulations. Further, with higher resolution PWM, the RtZ and common-mode circuits can be eliminated as long as the maximum output amplitude is limited, always ensuring at least one of the master clock period of the output waveform is zero.

Using PWM output also doubles the array elements and increases the quantizer resolution. When a 4-bit PWM is used, the OSR of the modulator can stay the same by cascading the PWM DAC elements. The higher resolution of the PWM output also allows for proper ±1 LSB TPDF dithering, linearizing the quantizer and eliminating any possible THD, noise floor modulation or cyclic behaviour of the ΔΣ modulators entirely [12].

The modulator could also be run at even higher rates, such as 4096 OSR. This high rate would allow for a 4-tap filter to be inserted, reducing the amount of noise-shaped RF content.

Further improvement can be made by optimizing the reconstruction filter implementation. The convolution engine can be modified when the sinc function amplitude is known, allowing to lower the convolution resolution as the sinc tails-off and only using the full precision near the main-lobe.

REFERENCES

[1] ARNALDI, Isacco. *Design of Sigma-Delta Converters in MATLAB®/Simulink®*. 2018. Caldogno, Italy: Springer Nature Customer Service Center, 2018, 282 pages. ISBN 3030082709, 9783030082703.

[2] KLEDROWETZ, Vilém. Nová struktura modulátoru delta-sigma nízkého řádu s vysokým rozlišením [online]. Brno, 2014 [q. 2019-11-16]. Available at: http://hdl.handle.net/11012/26759. Disertační práce. Vysoké učení technické v Brně. Fakulta elektrotechniky a komunikačních technologií. Ústav mikroelektroniky. Vedoucí práce Jiří Háze.

[3] Chord DAVE. *HI-FI NEWS* [online]. 2016, 2016(9), 4 [q. 2019-12-01]. Available at: https://www.chordelectronics.jp/static/pdf/news/dave/1603_HFN_Chord%20DAVE_hires.pdf

[4] SCHREIER, Richard and Gabor, TEMES. *Understanding Delta-Sigma Data Converters*. A JOHN WILEY & SONS. INC ., PUBLICATION.

[5] E.C. Dijkmans, B. Kup and J. Sneep. 1991. Bit stream D/A converter with 18-bit dynamic range. *AUDIO ENGINEERING SOCIETY*. Philips Research Labs, Eindhoven: Philips, 1991 February 19-22 (Presented at AUDIO the 90th Convention in Paris), 18.

[6] DUNN, Chris and Mark SANDLER. 1995. *Liearising Sigma-Delta Modulators using Dither and Chaos*. King's College, London: IEEE, p. 3.

[7] HOWARD, Keith. Tapping Into Better Digital Audio. *HIFI CRITIC* [online]. 2018 (JUL/AUG/SEP), 4. Available at: https://drive.google.com/file/d/1QhJo_Rh8sQsItFvJ1g-MnNlg9pBR6oja/view

[8] ECSE 4530 - Digital Signal Processing. *InfoCoBuild* [online]. 2014 Available at: http://www.infocobuild.com/education/audio-video-courses/electronics/ECSE4530-Fall2014-RPI/lecture-13.html

[9] *Auditory Perception* [online]. 2016. Available at: https://web.archive.org/web/20100410235208/http://www.cs.ucc.ie/~ianp/CS2511/HAP.html

[10] WANG, DeLiang and Guy J. BROWN. 2006. Computational Auditory Scene Analysis: Principles, Algorithms, and Applications. Wiley-IEEE Press.

[11] NENTWIG, Markus. Shared-multiplier polyphase FIR filter. *DSP Related* [online]. 2013 [2020-05-12]. Available at: https://www.dsprelated.com/showarticle/198.php

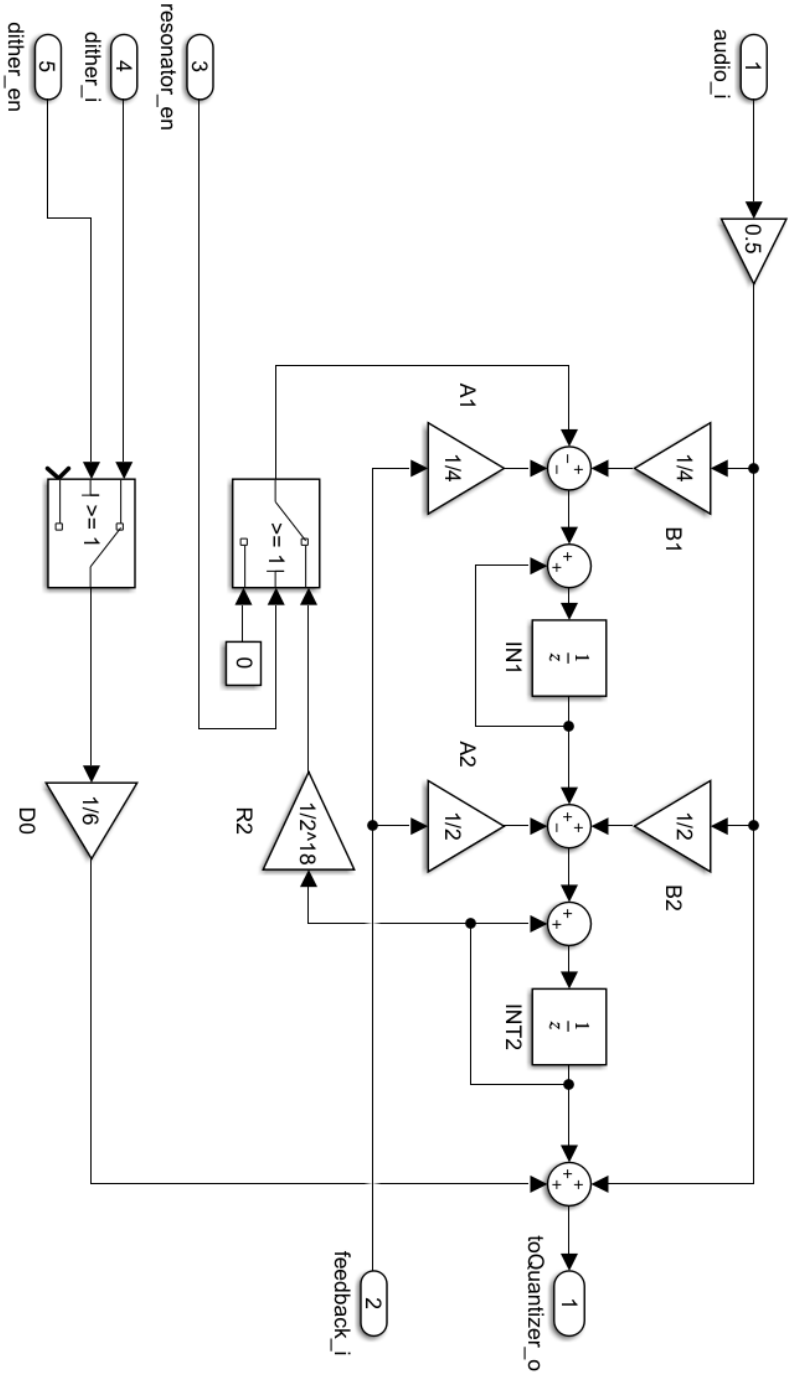
[12] LIPSHITZ, Stanley P. a John VANDERKOOY. Convention Paper 5395. *Audio Engineering Society*. Netherlands, 2001, (Presented at the 110th Convention), 11.

[13] DOHNAL, J. Návrh a vývoj programovatelného notch filtru pro audiotechniku .Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2018. 61s. Vedoucí bakalářské práce doc. Ing. Jiří Háze, Ph.D.

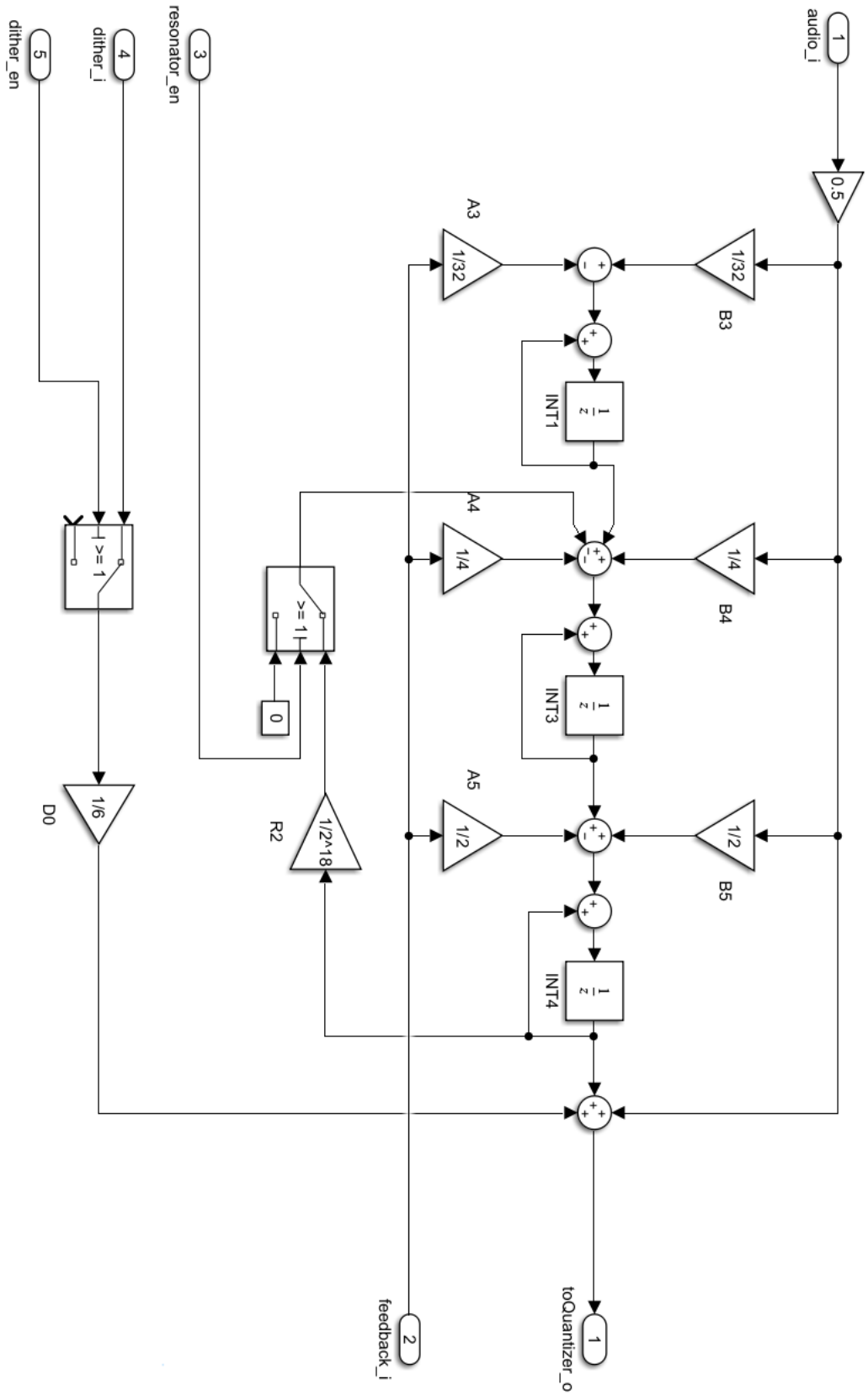
[14] *Python-deltasigma v0.2.2 documentation* [online]. [q. 2020-05-10]. Available at: https://python-deltasigma.readthedocs.io/en/latest/

A SIMULINK MODELS

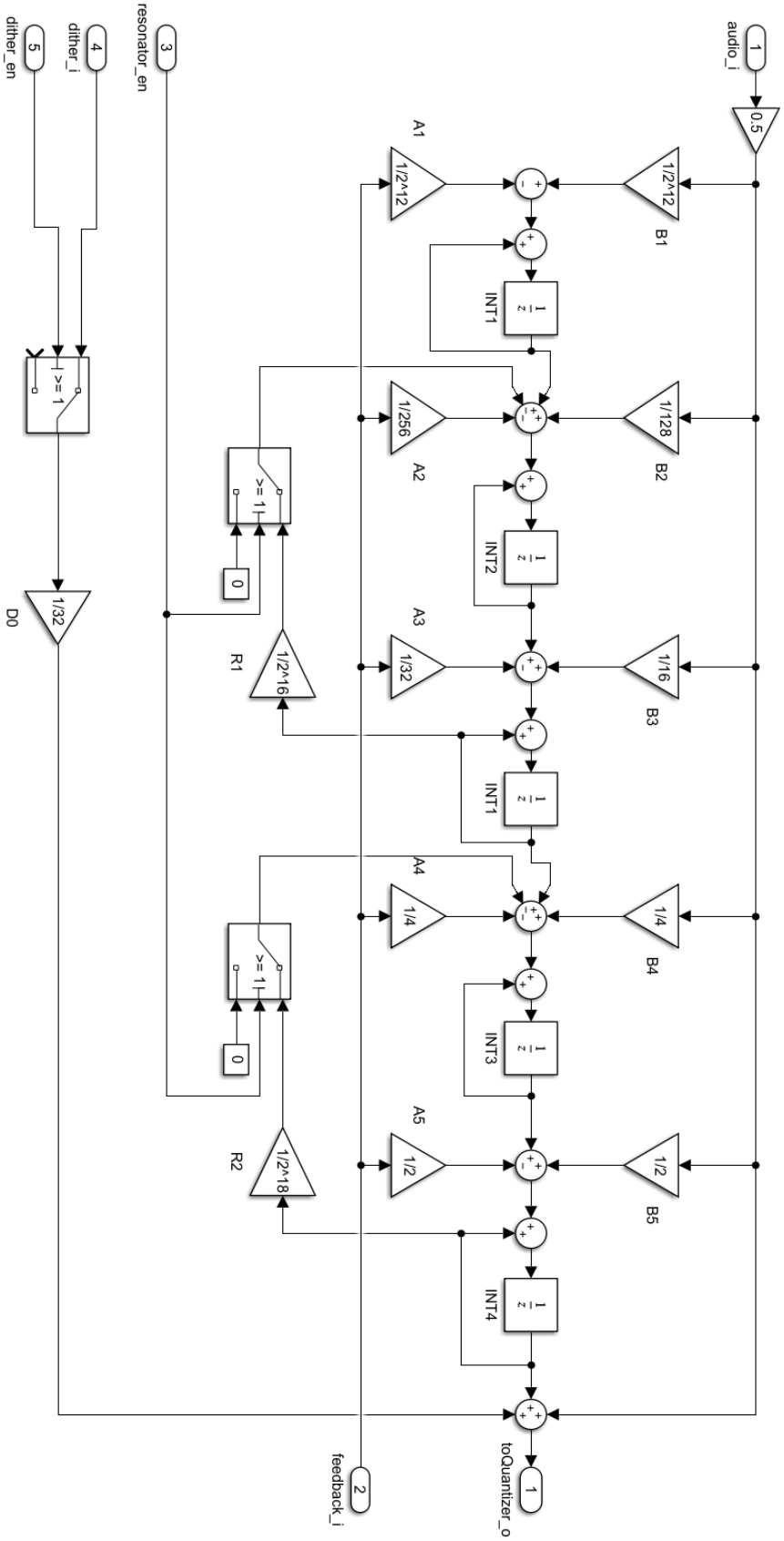
A.1 Second-order ΔΣ modulator loop



A.2 Third-order ΔΣ modulator loop

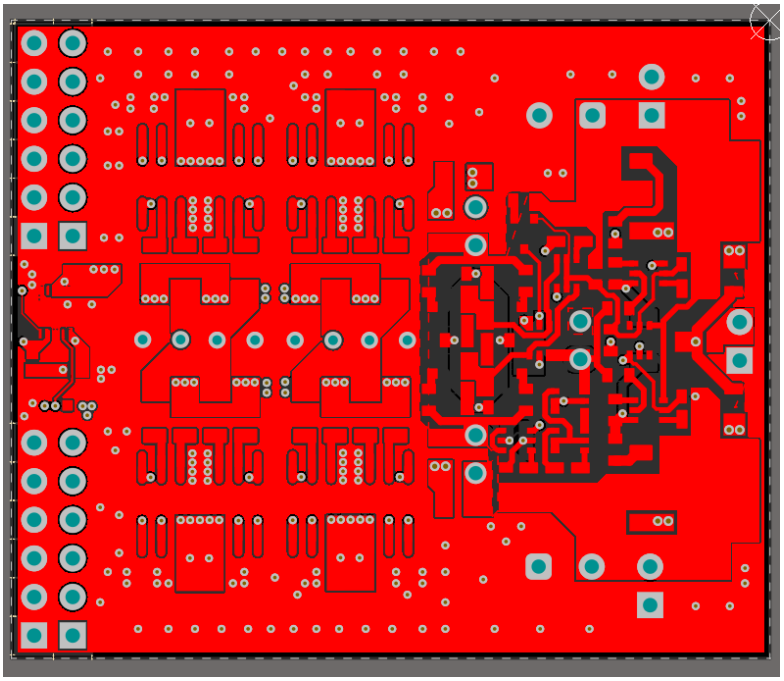


A.3 Fifth-order ΔΣ modulator loop

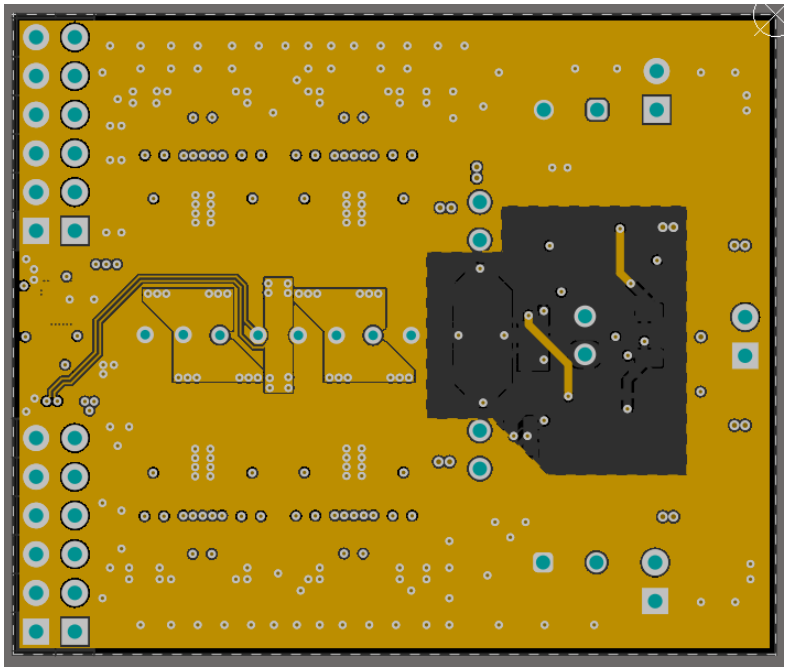


B DAC ARRAY

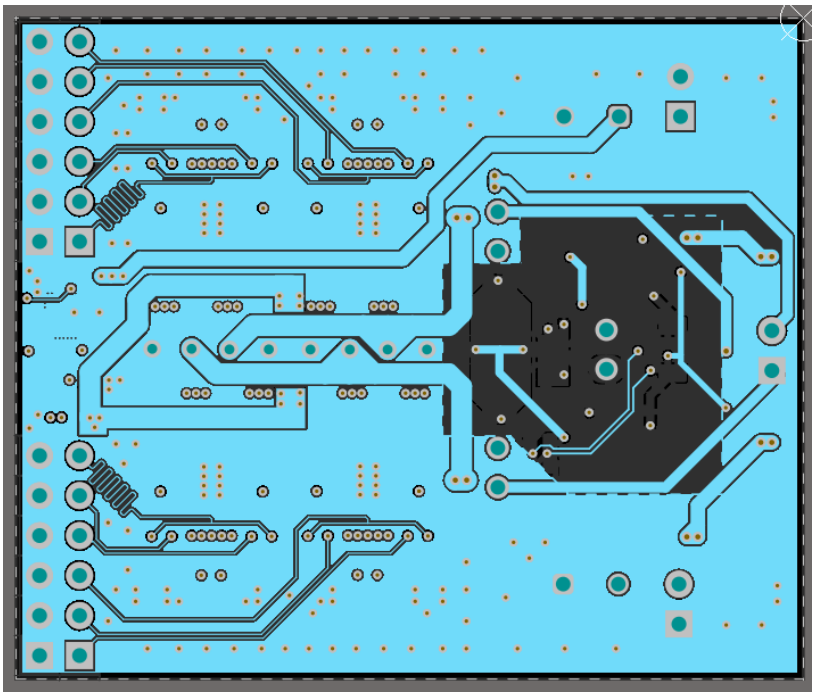
B.1 PCB layout



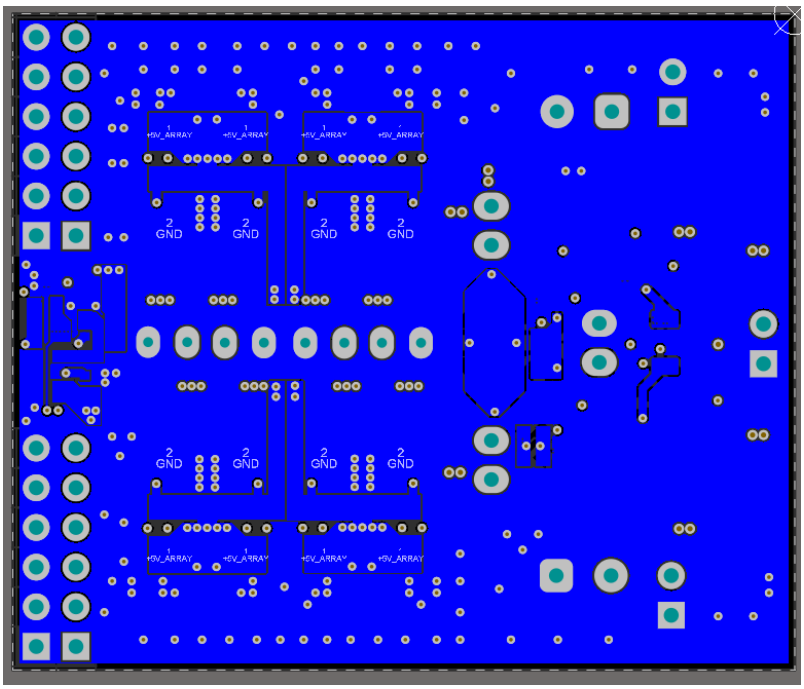
TOP layer of the DAC and output stage



IN2 layer of the DAC and output stage

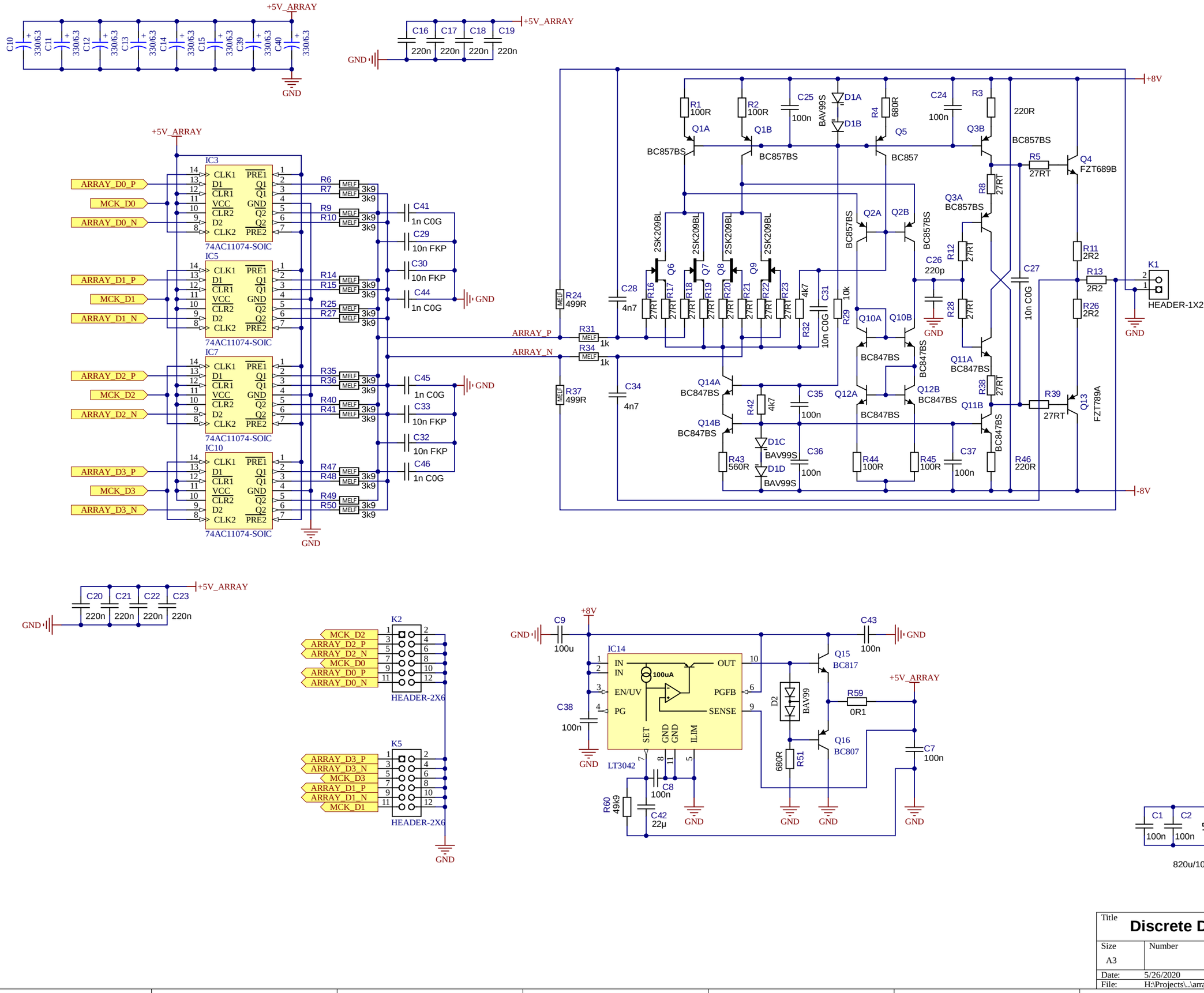


IN3 layer of the DAC and output stage



BOT layer of the DAC and output stage

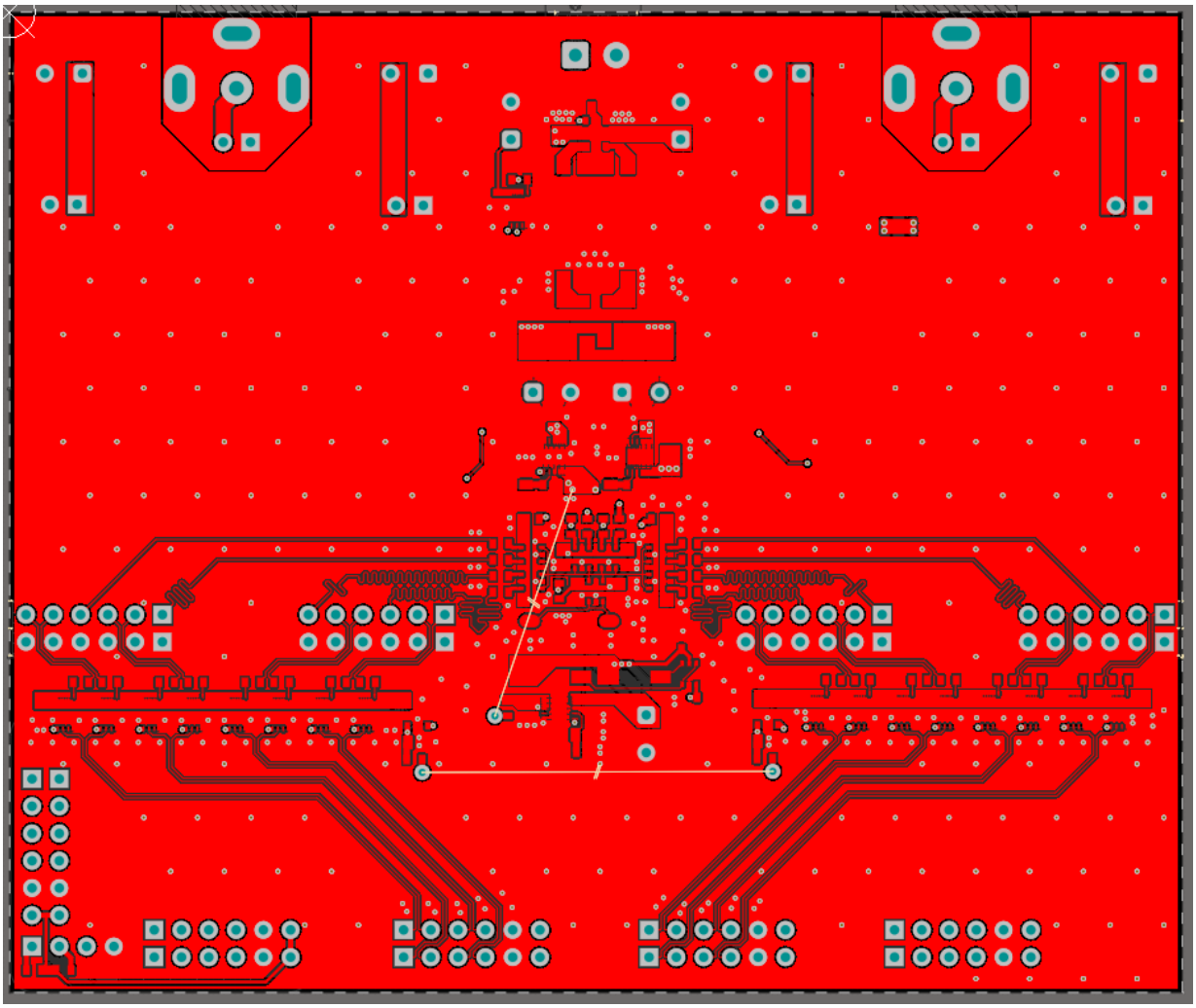
B.2 Schematic



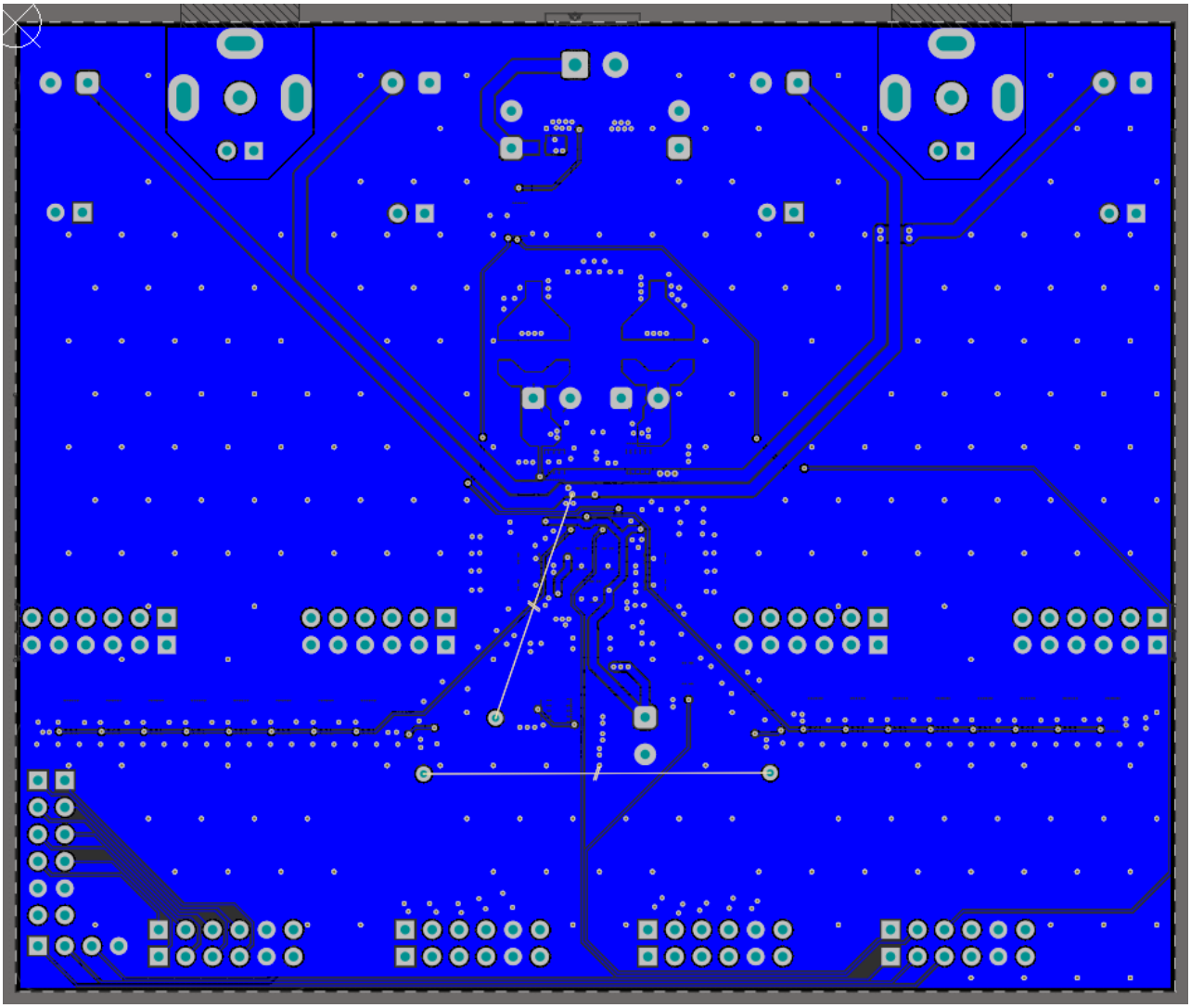
Title	**Discrete DAC**		**Array and OPA**
Size	A3	Number **C05-202-0001**	Revision **A**
Date:	5/26/2020		Sheet **1** of **1**
File:	H:\Projects\..\array.SchDoc		Drawn By:

C MOTHERBOARD

C.1 PCB layout

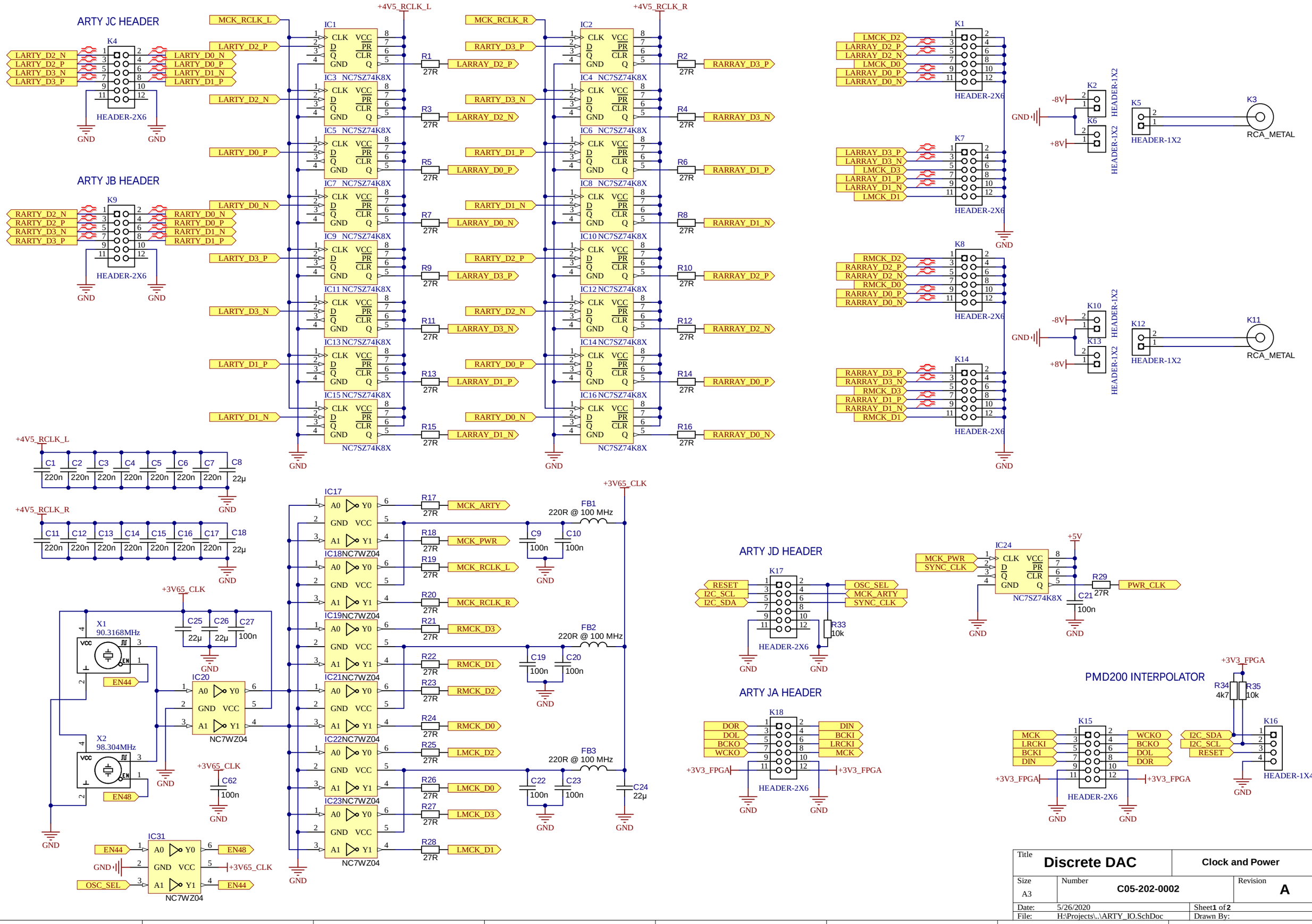


TOP layer of the interconnect motherboard

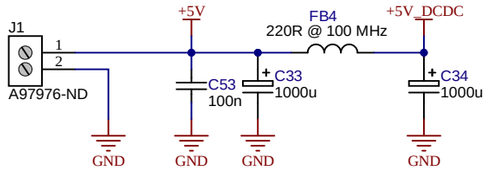


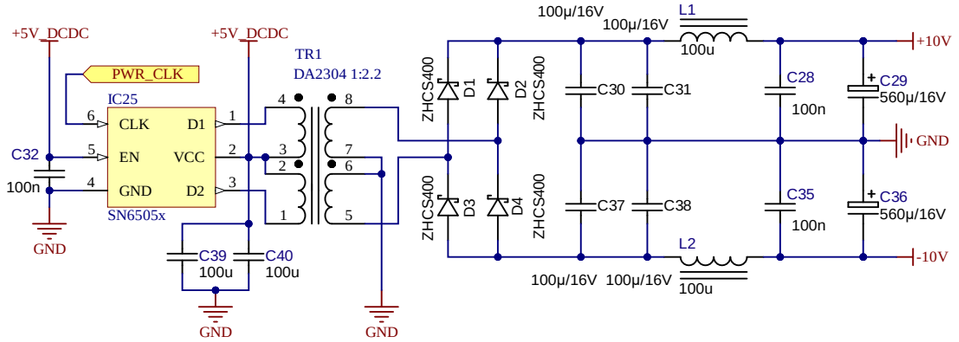
BOT layer of the interconnect motherboard

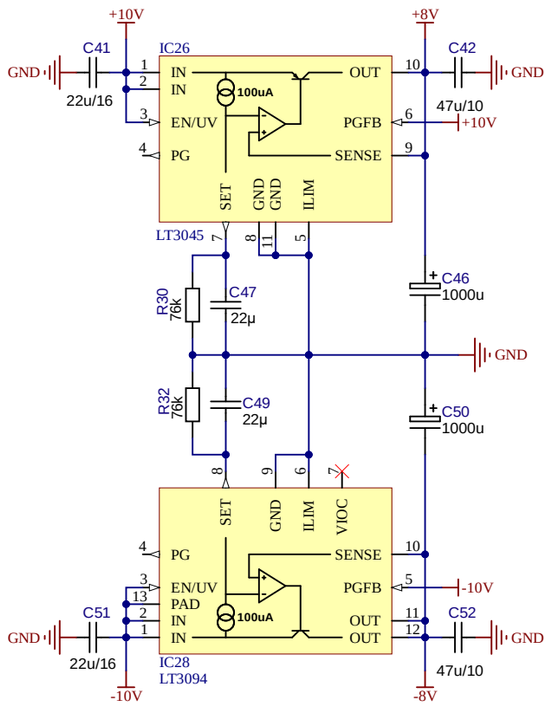
C.2 Schematic

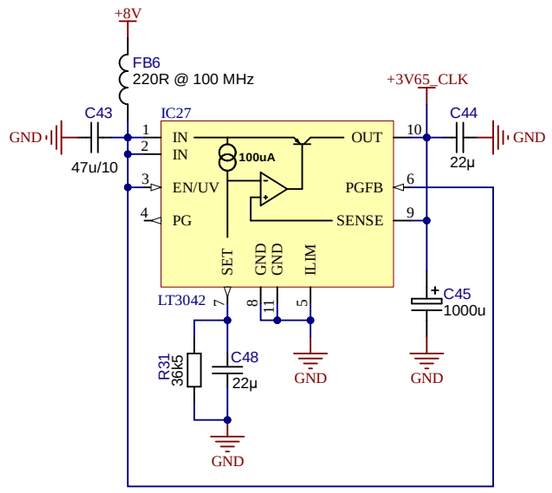


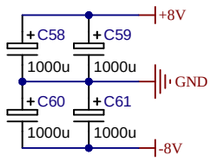
Title	**Discrete DAC**		**Clock and Power**
Size	A3	Number **C05-202-0002**	Revision **A**
Date:	5/26/2020		Sheet **1** of **2**
File:	H:\Projects\..\ARTY_IO.SchDoc		Drawn By:

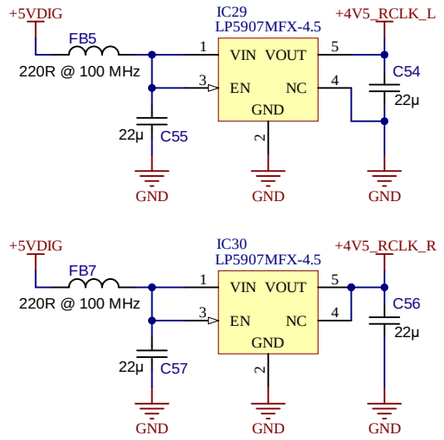












Title	**Discrete DAC**		**Clock and Power**
Size: A3	Number: **C05-202-0002**		Revision: **A**
Date: 5/26/2020		Sheet **2** of **2**	
File: H:\Projects\..\PWR.SchDoc		Drawn By:	