AN ANALOG IMPLEMENTATION OF THE PSYCHOACOUSTIC MODEL USED IN AUDIO COMPRESSION ALGORITHMS ON AN FPAA

A Dissertation
Presented to
The Academic Faculty

By

Lenno Liu

In Partial Fulfillment
of the Requirements for the Degree
Master of Science in the
School of Electrical and Computer Engineering

Georgia Institute of Technology

May 2023

AN ANALOG IMPLEMENTATION OF THE PSYCHOACOUSTIC MODEL USED IN AUDIO COMPRESSION ALGORITHMS ON AN FPAA

Thesis committee:

Dr. Jennifer Hasler
Department of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. David Anderson
Department of Electrical and Computer Engineering
Georgia Institute of Technology

Dr. Aaron Lanterman
Department of Electrical and Computer Engineering
Georgia Institute of Technology

Date approved: April 28, 2023

Close your eyes.

Count to one.

That’s how long forever feels.

Kurzgesagt

Dedicated to Pow Chu Moy, Warren Young, and Zachary Barth.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Hasler, for all of the great discussions and conversations had. Her intuition and insight in all manners of circuits were instrumental in the success of this project.

I would also like to thank the present and past members of the Integrated Computational Electronics (ICE) lab, for making this project possible in the first place: Pranav, Afolabi, Swagat, Praveen, Linhao, John, and Eric. Thank you for putting up with my constant barrage of ideas and data. Especially Pranav, who helped debug all sorts of bugs found in the RASP tools.

I would like to thank Pow Chu Moy (founder of HeadWize), Warren Young (admin of Tangentsoft), and Zachary Barth (founder of Zachtronics) for instilling my interest in analog electronics at a young age and for showing me the joy in engineering. Without you, I would not have found my calling.

I would like to thank my family for their continuous support through all the highs and lows of my life: Angus, Mom, Dad, and Nathan. Thank you for all the understanding, acceptance, and unconditional love that help make me who I am today.

Lastly but not least, I would like to thank the committee members, Dr. Anderson and Dr. Lanterman, for their discussions and feedback on this project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Lossy audio compression takes advantage of how humans perceive sound to achieve high compression ratios. This is done by analyzing the audio to determine what is audible to the human ear using a psychoacoustic model, and preferentially allocating more bits to encode the audible portions while cutting down on the bits used to encode the masked or inaudible portions. This method allows for an approximately tenfold reduction in the memory required without a perceptible loss in quality, and is used everywhere in the modern world from VoIP calls to music storage and streaming services. All current audio compression algorithms work on a digital in, digital out basis converting from a pulse-code modulated (PCM) signal to a bitstream output. This work aims to provide a basis to create a mixed-signal encoder that takes an analog input for the same digital output, by implementing an analog psychoacoustic model on a field-programmable analog array (FPAA).

CHAPTER 1

PSYCHOACOUSTICS AND LOSSY AUDIO COMPRESSION

1.1 The Equal Loudness Contour

Sound waves in air are oscillations in atmospheric pressure, and are characterized by their frequency and amplitude. The frequency range of interest for human hearing is generally accepted to be between 20 Hz to 20 kHz. The sensitivity of the human ear changes across this range, being most sensitive to frequencies near 3 kHz and less sensitive to frequencies near the limits.

To characterize the human ear, the most common reference tone used is a sinusoidal signal of frequency 1 kHz at 40 dB SPL. Sound pressure level (SPL) refers to a measurement of the amplitude of the pressure wave that is referenced to the threshold of human hearing at 1 kHz:

$$SPL = 20 \log_{10} \left(\frac{p}{p_{ref}} \right)$$

$$p_{ref} = 2 \times 10^{-5} \, \mathrm{Pa}$$

Due to the uneven frequency response of the human ear, a 40 dB SPL signal at 100 Hz is not perceived to be equally loud to a 40 dB SPL signal at 1 kHz. Instead, a higher SPL is required to obtain the same perceived loudness. The SPL required to obtain the same perceived loudness as the 40 dB SPL signal at 1 kHz is known as the 40-phon equal loudness contour, shown in Figure 1.1 [1]. To obtain the same perceived loudness at 100 Hz, an SPL of $\sim$65 dB is needed. Note that the phon is a unit of perceived loudness empirically derived from psychoacoustic studies, referenced to a 1 kHz test tone with the equivalent phon number in SPL.

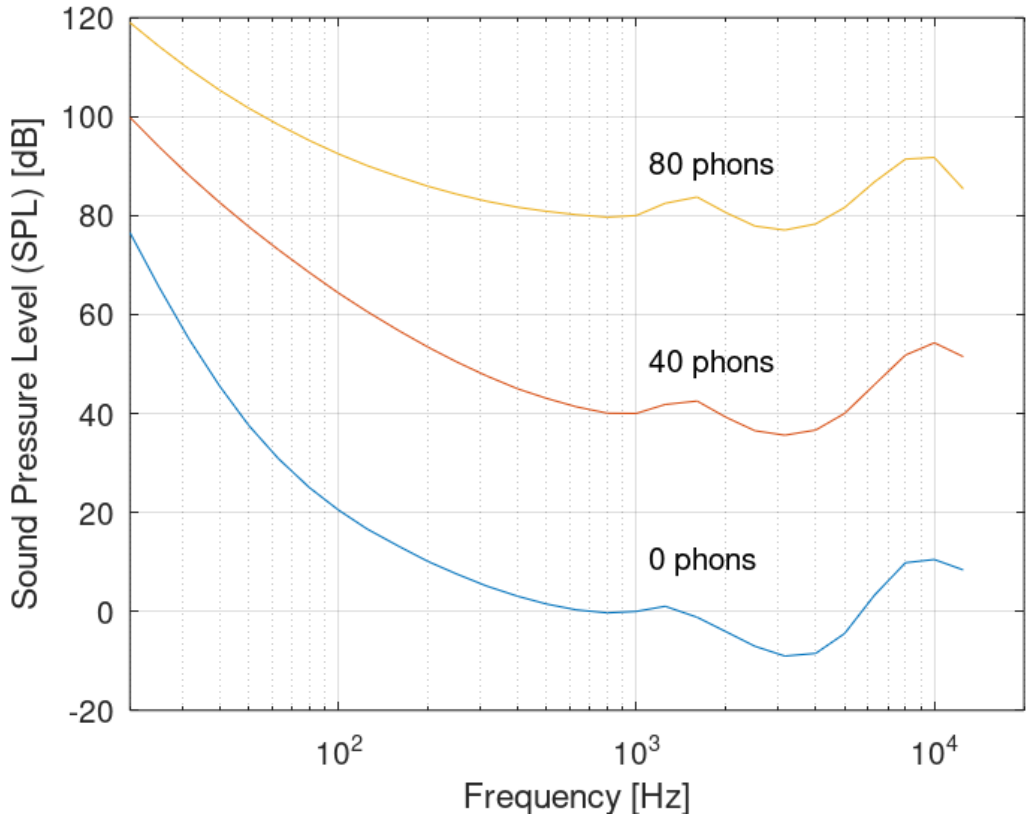

Figure 1.1: A selection of equal loudness contours at three different phon levels. Note that the 0 phon contour is also known as the absolute threshold of hearing.

Around 40 phons and above, the perceived loudness scales roughly linearly to SPL for any fixed frequency. This indicates that for a linear increase in perceived loudness, an exponential increase in pressure is needed. This means that human hearing spans roughly 6 decades of dynamic range in pressure between the absolute threshold of hearing and the threshold of pain at $\sim$120 dB [2].

1.2 Frequency and Perceived Pitch

Just like how loudness is a perceived measure of a sound wave's pressure, pitch is the perceived measure of a sound wave's frequency. A higher frequency tone is perceived as a higher pitch. Figure 1.2 shows the relation between the notes on a piano and the fundamental frequencies associated with each note.

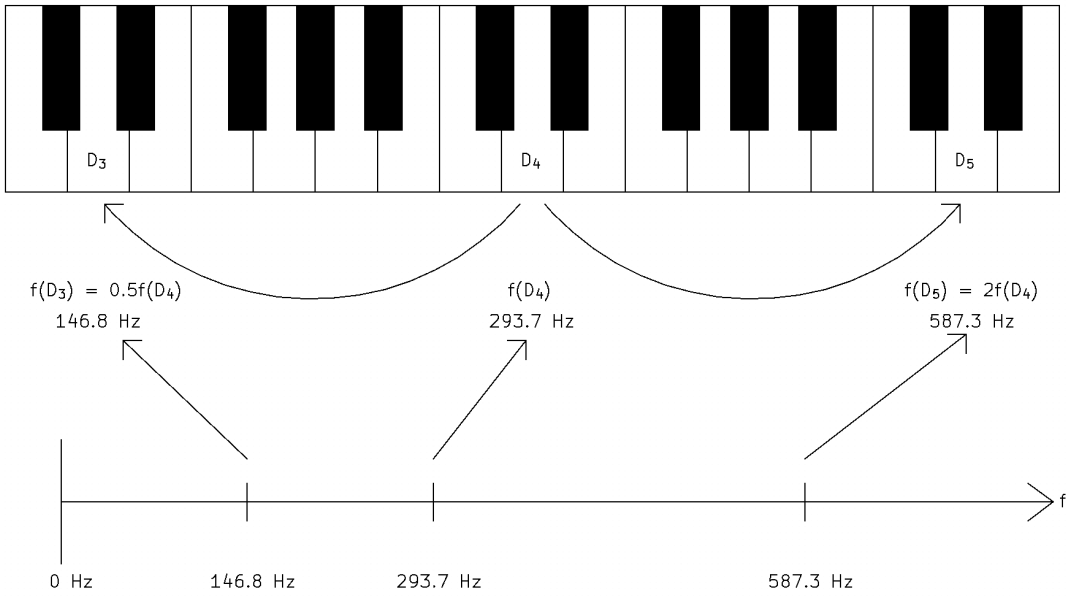

Figure 1.2: Correlation between frequency and perceived pitch. Each octave higher on a piano corresponds with a doubling in frequency, and vice versa.

In western music, the most fundamental interval between two pitches is the octave. The octave is divided into twelve equal steps called semitones. Going up an octave in pitch on the piano corresponds with a doubling in frequency, while going up one semitone corresponds with a multiplication of $\sqrt[12]{2}$ in frequency. This observed exponential relationship between perceived pitch and frequency holds across the entire range of human hearing from 20 Hz to 20 kHz, or approximately 10 octaves. The human ear is highly sensitive to changes in pitch; the typical adult is able to easily recognize pitch differences of $1/4$ semitones with no musical training, corresponding to a $\sim 1.45\%$ change in frequency [3].

1.3 Auditory Masking

Auditory masking is a psychoacoustic phenomenon where the frequency response of the human ear adapts in response to an input signal. There are two types of auditory masking: simultaneous masking and temporal masking [4].

Simultaneous masking, also known as spectral masking, is the phenomenon where a

large input signal suppresses the ear's ability to sense a smaller signal nearby in frequency. Figure 1.3 shows an example of simultaneous masking where the presence of a strong 1 kHz signal renders a weak signal at 1.2 kHz inaudible.

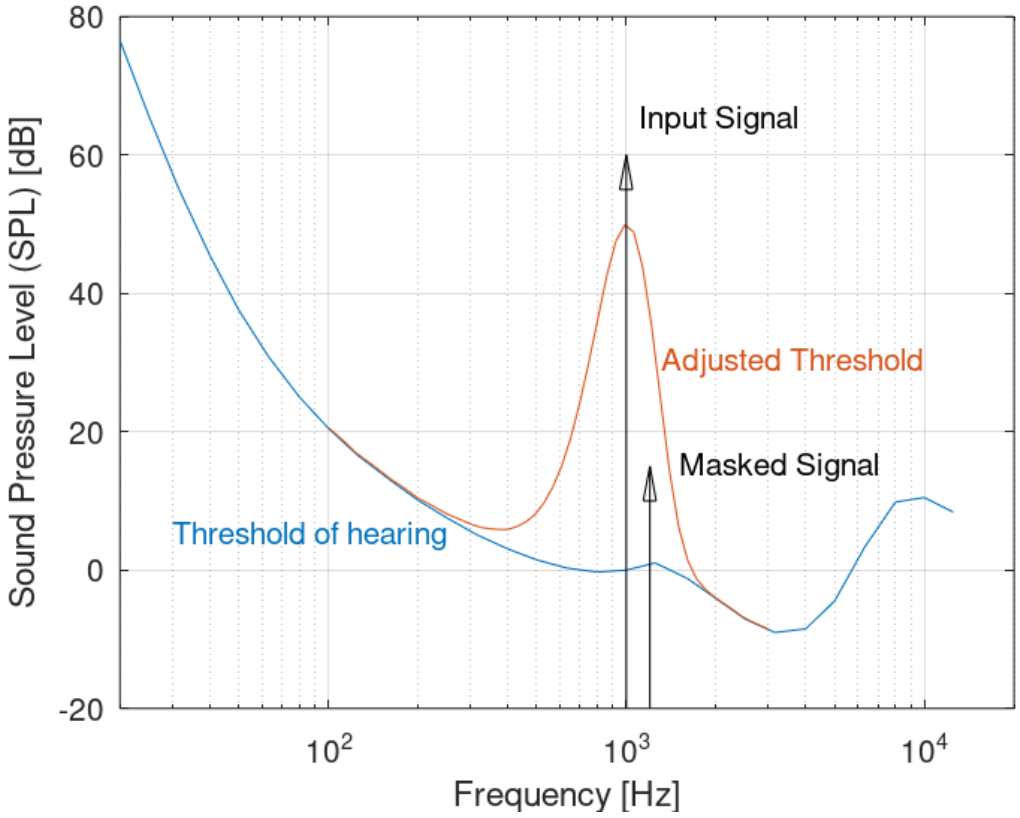

Figure 1.3: An example of simultaneous auditory masking. The large input signal prevents small signals nearby in frequency from being audible.

Temporal masking is the phenomenon where an input signal suppresses the ear's ability to sense other signals occurring immediately before or after the signal. This effect is generally not incorporated into the psychoacoustic models of perceptual audio encoders.

1.4 The MP3 Encoder

Figure 1.4 shows the block diagram of the MP3 encoding process [5].

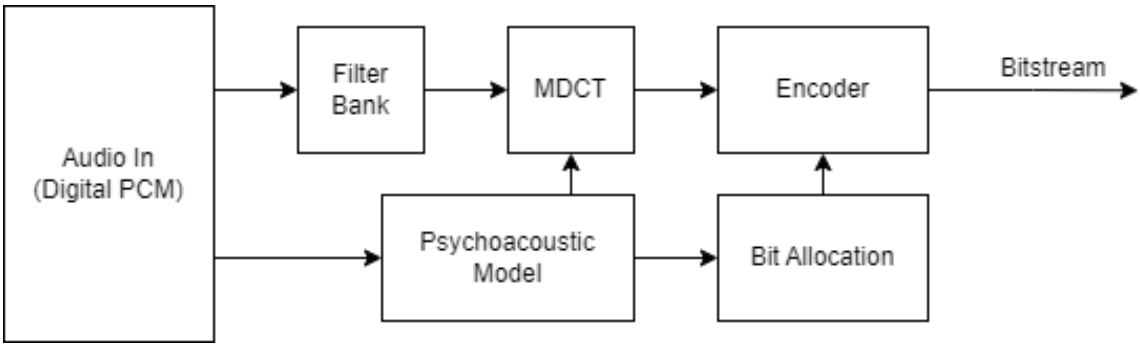

Figure 1.4: A flowchart detailing the MP3 encoding process. The quantization, Huffman coding, and bitstream formatting processes have been simplified into the 'Encoder' block.

The audio input is first split up into frames that are $\sim$26 ms in length, each containing 1152 samples. These frames are passed through a 32-parallel, linearly spaced filter bank to analyze the audio's spectral content. The filter bank is linearly spaced from 20 Hz to 20 kHz, giving each filter a bandwidth of $\sim$650 Hz. Each subband is then passed through a discrete cosine transform designed to operate on overlapping frames, known as a modified discrete cosine transform (MDCT). Simultaneously, the input audio is passed through an FFT into the psychoacoustic model to compute the masking effects, which are used to scale the coefficients used in the MDCT. The psychoacoustic model is also used to determine the number of bits allocated to each subband in the encoding process. The encoder block shown in Figure 1.4 contains the quantization, Huffman coding, and bitstream formatting processes required to obtain the final MP3 output.

1.5 Mixed-Signal Encoder

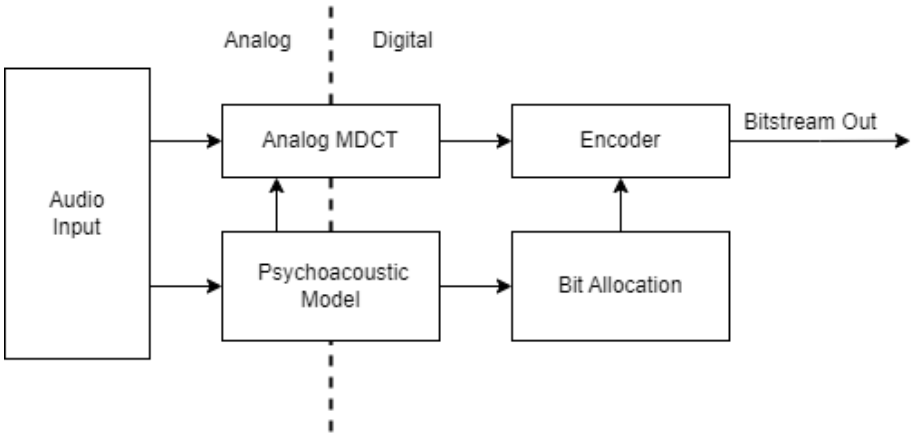

Figure 1.5: Block diagram for a proposed mixed-signal encoder. The quantization, Huffman coding, and bitstream formatting processes have been simplified into the 'Encoder' block.

Figure 1.5 shows the block diagram for a proposed mixed-signal encoder. In this encoder, the audio input is an analog input instead of a digital PCM signal. This has the effect of allowing the FFT and digital psychoacoustic model block to be changed to an exponentially spaced filterbank and analog psychoacoustic model. Additionally, an analog implementation of the MDCT may be used to process the subband signals before being converted to a digital signal for quantization in the encoder [6] [7].

This method of encoding audio directly from an analog input to a digital bitstream can be much more power efficient than the equivalent digital implementation [8] [9], and may be interfaced directly to the output of a microphone. This has potential applications where audio encoding is required at low power, such as battery powered devices, mobile phones, and other systems that can benefit from reduced power consumption.

CHAPTER 2

THE FIELD PROGRAMMABLE ANALOG ARRAY

Figure 2.1: The FPAA device used: version 3.0A, board #12

This chapter describes the reconfigurable analog circuit platform, the FPAA, that is used as a basis for building the analog psychoacoustic model. Figure 2.1 is a photo of the specific FPAA device used, designated as version 3.0A board #12.

2.1 Floating Gate MOSFETs

The floating gate MOSFET is the basic building block of the FPAA. The difference between this device and the normal MOSFET is at the gate terminal, which is entirely insulated. This causes a fixed charge to stay on the gate, which alters the bias point of the device and allows for programmability [10]. Figure 2.2 shows the simplified symbol of the FG PFET and its expanded equivalent. During operation, the programming MOSFET is disconnected via

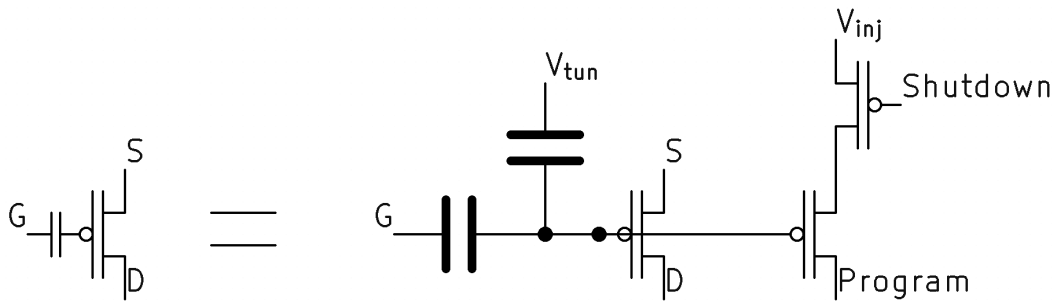

Figure 2.2: The floating-gate PFET. This device forms the basis of programmable analog circuits on the FPAA.

the shutdown MOSFET and only the device MOSFET is active.

2.1.1 Programming Floating Gate MOSFETs

The charge on a FG MOSFET can be precisely controlled by changing the number of electrons on the floating gate node. This is done by Fowler-Nordheim tunneling and hot carrier injection. In Fowler-Nordheim tunneling, the potential barrier of the insulator is effectively reduced in thickness under high electric fields, causing an exponential increase in the probability of quantum tunneling of electrons out of the gate. In hot carrier injection, the high electric field gives electrons additional kinetic energy as it conducts through the MOSFET and exponentially increases the probability of the electron tunneling through the insulator and into the gate. These two effects are exploited to change the amount of charge present on any floating gate in the FPAA. To measure the state of the FG MOSFET, the current coming out of the program pin of the MOSFET is measured. As such, it is standard practice to program FG MOSFETs to a target current value instead of a target gate voltage value.

2.2 FPAA Circuit Elements

The FPAA contains an array of computational analog blocks (CABs), each of which contains an assortment of analog circuit components that can be connected to each other to create analog systems. Table 2.1 catalogs the circuit elements available inside a single CAB.

Table 2.1: Circuit Elements Available in One CAB

Circuit Element	Number Available
OTA	2
FGOTA	2
Capacitor Bank	4
T-Gate	4
NFET	2
PFET	2
N-mirror	2
Routing FG PFET	990

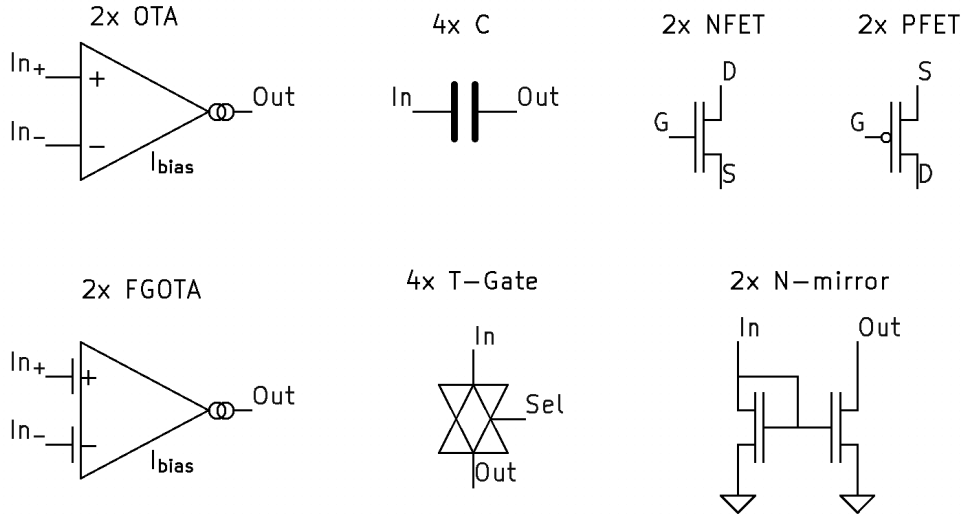

Figure 2.3: The circuit elements available in one CAB. Each of these components can be wired to each other using FG PFETs to create various analog circuits.

The routing FG PFETs exist in a series of crossbar arrays inside the CAB that allow the connection of circuit elements to each other. Each FG PFET can be programmed to

a ‘C’ low-impedance connection, a ‘T’ target current connection, or a ‘-’ disconnected connection.

The bias current to each OTA and FGOTA is supplied by a FG PFET and can be programmed to a set current. This allows for the fine tuning of filters created with such devices. The FGOTA is simply an OTA with FG PFETs instead of normal PFETs as inputs, which extends the linear range of the OTA and also allows for the programming of offset voltages into the device. The transconductance of these devices can be calculated as

$$g_{m,ota} = i_{out}/v_{in},$$

$$i_{out} = I_{bias} \tanh\left(\frac{V_+ - V_-}{2U_T}\right).$$

Within the linear region of the $\tanh$ function, this can be simplified to

$$i_{out} = I_{bias} v_{in}/2U_T,$$

$$g_{m,ota} = \frac{I_{bias}}{2U_T}.$$

The transconductance of the FGOTA is slightly different due to the capacitive divider on the input. Assuming operation in the expanded linear region of the $\tanh$ function, it can be simplified to be

$$g_{m,fgota} = \frac{C_{in}}{C_T}\frac{I_{bias}}{2U_T},$$

where C_{in}/C_T is the ratio of the input capacitance to the total gate capacitance of the floating gate input FETs to the FGOTA.

Each capacitor bank contains three capacitors in parallel, designed to be 64 fF, 128 fF, and 256 fF. This allows a selection of 7 different capacitance values for each bank ranging from 64 fF to 448 fF. Note that parasitic capacitances inside the CAB and routing is around 160 fF for each local line and routing outside the CAB, and may have a large influence on

circuit operation [11].

2.3 FPAA Tool Chain

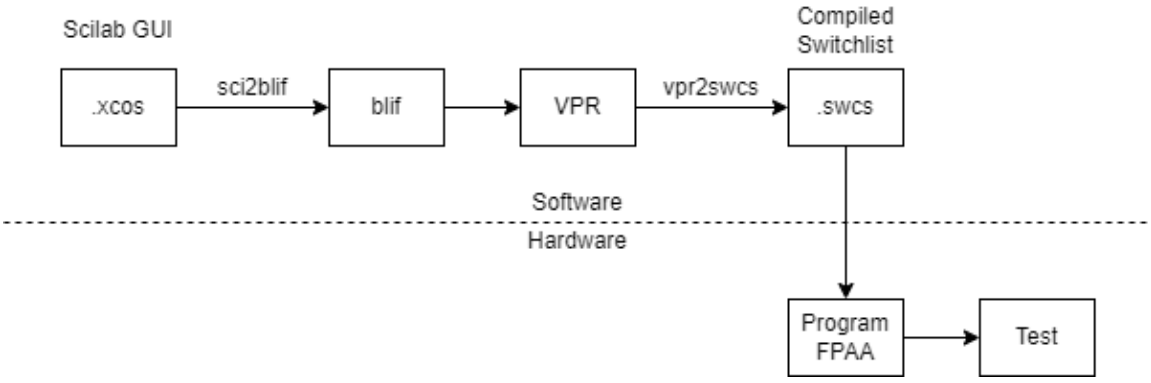

Figure 2.4: The FPAA tool chain. Circuits are first created via the Scilab GUI, then compiled and programmed onto the FPAA.

Figure 2.4 shows the workflow process for creating a system on the FPAA [12]. Creating a physical system begins with the open source RASP tools on Scilab that runs in an Ubuntu VM. The desired system is first created in Xcos, a graphical editor similar to Simulink that interfaces with Scilab. Once the Xcos file is created, the design is compiled by the RASP tools to create a switchlist. This switchlist is then programmed into the FPAA, and run by the FPAA after a command is issued from the RASP tools.

CHAPTER 3

DESIGN OF THE ANALOG PSYCHOACOUSTIC MODEL

The analog psychoacoustic system was designed top-down and broken into three analog blocks and one mixed-signal block, compiled into two unique CABs and a variation on the second CAB. These four blocks consist of the C^4 bandpass filter block, min detect block, DSWTA block, and I2F block, shown in Figure 3.1. The first two blocks fit into one CAB, and the latter two blocks fit into another CAB. This basic four-block system was then parallelized and tuned to obtain the final system. An 8-parallel system tuned in the range of 300 Hz to 3 kHz, also known as the telephone band [2], was chosen as it is directly applicable to applications in telecommunications. It also has approximately the same filterbank spacing of 8 / decade compared to 10.67 / decade in a 32-parallel system with a range from 20 Hz to 20 kHz.

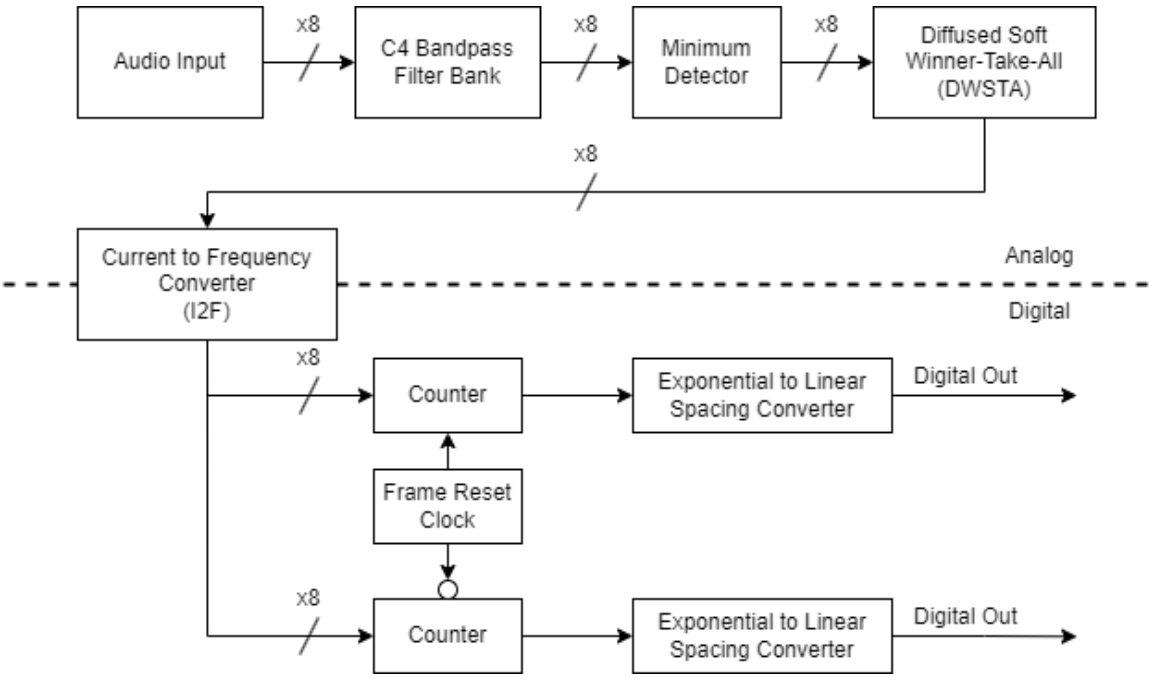

Figure 3.1: Block diagram of the 8-parallel psychoacoustic model. Implementation on the FPAA is limited to the output of the I2F converter.

3.1 The C^4 + Min Detect CAB

3.1.1 C^4 Bandpass Filter

The C^4 bandpass filter block is a low power, tunable bandpass filter implemented on the FPAA [13]. It consists of two OTAs feeding into each other, an input capacitor C_{in}, a feedthrough capacitor C_2, and an output capacitor C_{out}, shown in Figure 3.2. This circuit's transfer function can be derived to be

$$H(s) = -\frac{C_{in}}{C_2}\frac{s\tau_l(1-s\tau_f)}{1+s(\tau_l+\tau_f(\frac{C_{out}}{C_2}-1))+s^2\tau_h\tau_l},$$

where

$$\tau_f = C_2/g_{mgain},$$

$$\tau_l = C_2/g_{mfb},$$

$$\tau_h = \frac{C_T C_{out} - C_2^2}{C_2 g_{mgain}},$$

and C_T is the total capacitance on the V_{in-} node.

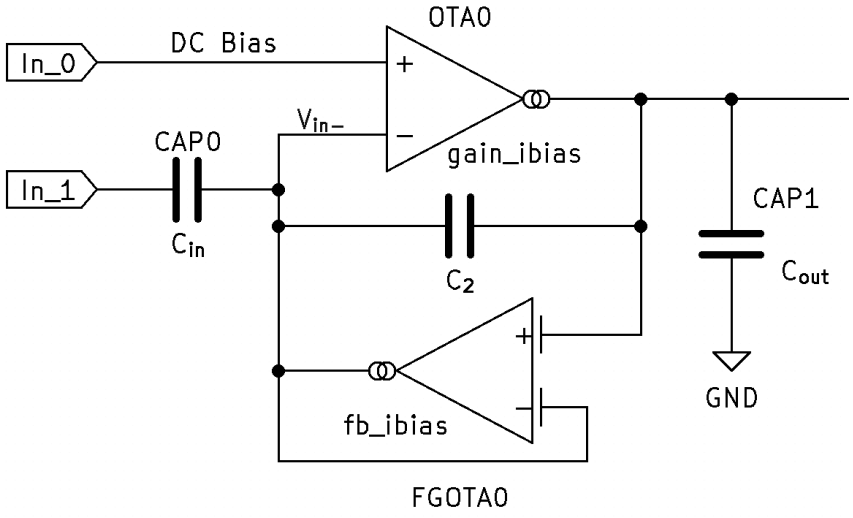

Figure 3.2: The C^4 bandpass filter as compiled on an FPAA. Note that the frequency response is a function of four tunable variables: input capacitance, output capacitance, gain bias current, and feedback bias current.

In practice, C_2 is a parasitic capacitance that is kept small to reduce the bias current requirements needed to obtain the desired gain and center frequency for each filter programmed in the C^4 filter bank.

Barring the parasitic capacitance C_2, the C^4 filter block has four variables that can be tuned to achieve the desired frequency response. These are C_{in}, C_{out}, I_{bias} of the gain OTA (gain_ibias), and I_{bias} of the feedback OTA (fb_ibias). Note that a FGOTA is used for the feedback to achieve a greater linear range, and that the positive input of the gain OTA is used to set the DC level of the output.

Table 3.1: Tuned C^4 Filter Bank Parameters

CAB Location	C_{in}	C_{out}	gain_ibias	fb_ibias	Target f_c	Achieved f_c	Gain	Temperature-adjusted f_c
[1;3;1]	2	1	5.7 nA	3.6 nA	2600 Hz	2819 Hz	5.49 dB	2502 Hz
[1;3;2]	2	1	4.1 nA	2.6 nA	1950 Hz	2152 Hz	5.82 dB	1991 Hz
[1;3;3]	2	1	4.3 nA	2.2 nA	1460 Hz	1601 Hz	5.47 dB	1422 Hz
[1;3;4]	2	1	3.0 nA	1.6 nA	1100 Hz	1194 Hz	5.74 dB	1060 Hz
[1;3;5]	2	2	2.4 nA	1.9 nA	822 Hz	921 Hz	5.17 dB	818 Hz
[1;3;6]	2	5	1.6 nA	1.3 nA	616 Hz	726 Hz	5.50 dB	645 Hz
[1;3;7]	2	5	1.5 nA	1.2 nA	462 Hz	516 Hz	5.95 dB	458 Hz
[1;3;8]	2	6	1.3 nA	1.1 nA	345 Hz	421 Hz	5.40 dB	374 Hz

Table 3.1 records the parameters used to create the C^4 filter bank of eight exponentially-spaced bandpass filters between 300 Hz and 3 kHz. The frequency response of each filter was characterized between 100 Hz and 10 kHz with a V_{pp} = 0.1 V input signal using a NI myDAQ, shown in Figure 3.3.

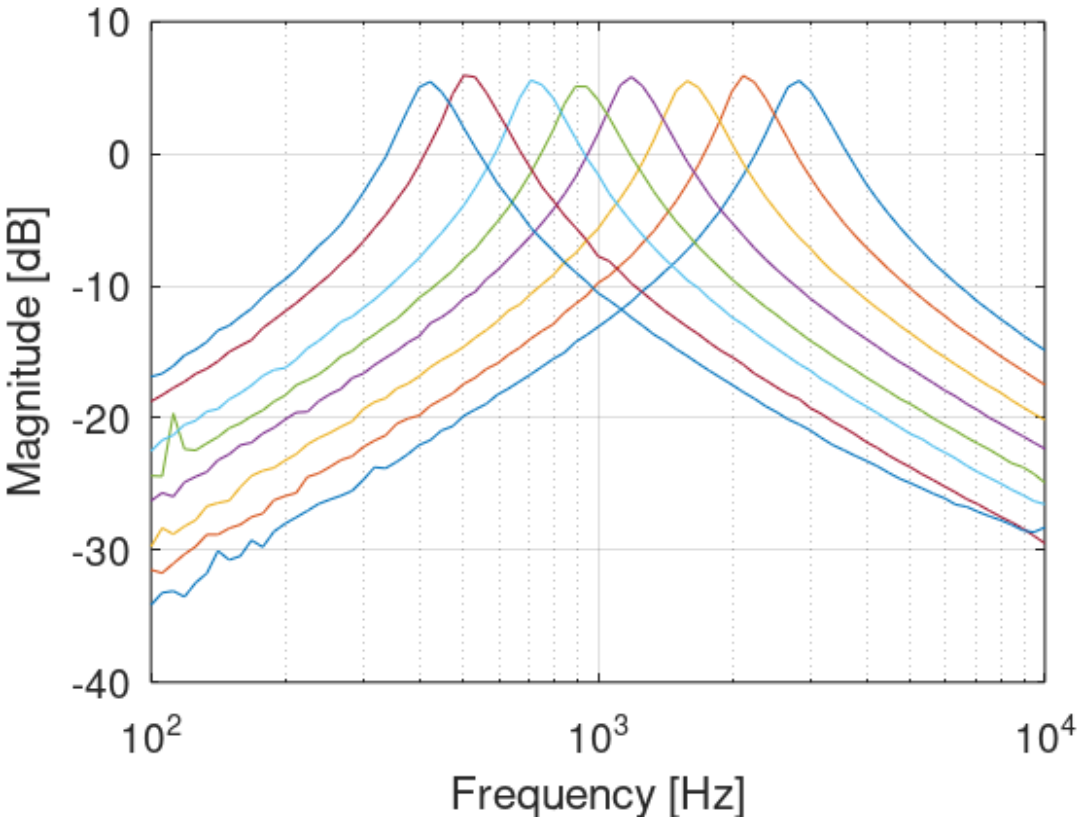

Figure 3.3: Frequency response of a tuned 8-channel filter bank to a V_{pp} = 0.1 V input signal. Each C^4 filter was compiled and programmed to set CABs, and manually tuned by adjusting capacitances and bias currents. The tuning compensates for mismatch between CABs to achieve a precise center frequency, gain, and Q factor for each filter.

Temperature Compensation

The C^4 bandpass filter bank as constructed has a large amount of temperature dependence due to the biasing structure of the OTAs. Each OTA is biased by a FG PFET operating in the subthreshold regime, giving an I-V relationship of

$$I_{ds} = I_s e^{\frac{\kappa V_g - V_s}{2U_T}},$$

where

$$U_T = \frac{k_B T}{q} \approx 25.8\,\text{mV@300}\,\text{K}.$$

If the numerator of $\kappa V_g - V_s$ is taken to be 0.5 V, a small change of 2 K results in a current change of

$$\frac{I_{ds1}}{I_{ds2}} = \frac{e^{\frac{0.5}{2U_{T300K}}}}{e^{\frac{0.5}{2U_{T298K}}}} \approx 0.937,$$

or about 6.3%. Since the transfer function of the C^4 bandpass function is dependent on the transconductance of the OTAs used, this causes the time constants τ_f, τ_l, and τ_h to shift proportional to $1/I_{bias}$. This means that the corner frequency will also shift by roughly 6.3% per 2 K.

If the discrepancy between the achieved f_c and target f_c is assumed to be solely attributed to temperature shift, then a calculated temperature shift of $\sim$4 K has occurred between programming the desired f_c and measuring the resulting f_c.

To greatly reduce the temperature dependence of the C^4 bandpass filter, one may use a floating gate bootstrap circuit [14]. Figure 3.4 shows the floating gate bootstrap circuit that generates a V_{out} that compensates for the temperature depedence of FG PFETs used in biasing the OTAs on the FPAA.

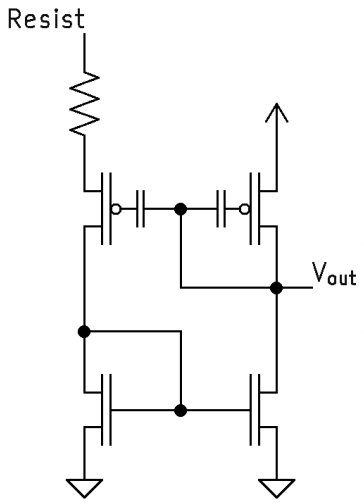

Figure 3.4: The floating gate bootstrap circuit. This circuit allows for the generation of a V_{out} compensated for temperature, greatly reducing temperature-induced current drift.

3.1.2 Minimum Detector

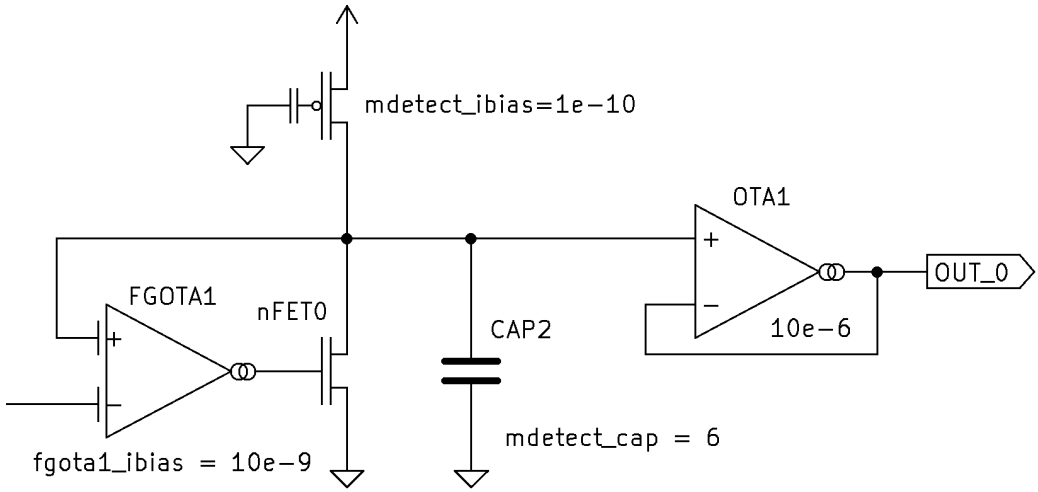

Figure 3.5: The minimum detector circuit as compiled on an FPAA.

The min detect block is used as an envelope detector for the output of the C^4 bandpass filter block. It is also used to tune any DC offsets introduced before the DSWTA. This is done by changing the charge at the inputs of the FGOTA. Figure 3.5 shows the schematic of the min detect block [15].

The FGOTA is used in a negative feedback configuration. As the input voltage at V_- of the FGOTA decreases below V_+, the output of the FGOTA increases, causing current to flow through nFET0. This decreases V_+ until it reaches V_-.

When $V_- > V_+$, the output of the FGOTA decreases to near GND. This causes nFET0 to turn off, and the voltage at V_+ rises slowly as current flows through the FG PFET. This rate is

$$\frac{dV_+}{dt} = \frac{I_{FGPFET}}{C_T},$$

where C_T is the total capacitance on the V_+ node. Since the routing of the V_+ node is within the CAB, it can be estimated to have a parasitic capacitance of $\sim$160 fF for a total capacitance of $\sim$550 fF [11]. This gives an estimated $\frac{dV_+}{dt}$ of $\sim$180 mV/ms. In practice, the slew rate is observed to be much lower at $\sim$16 mV/ms due primarily to a much smaller

I_{FGPFET} caused by how the FPAA programs bias currents less than 1 nA.

3.1.3 CAB Behavior

The C^4 bandpass filter block and min detect block were combined into a single CAB, and multiple copies of this CAB were programmed onto the FPAA with each CAB tuned to a different corner frequency as specified in Table 3.1. Figure 3.6 and Figure 3.7 show the combined response of four C^4 + min detect CABs for locations [1;3;4] to [1;3;7] to a V_{pp} = 0.2 V and V_{pp} = 0.5 V input, respectively.

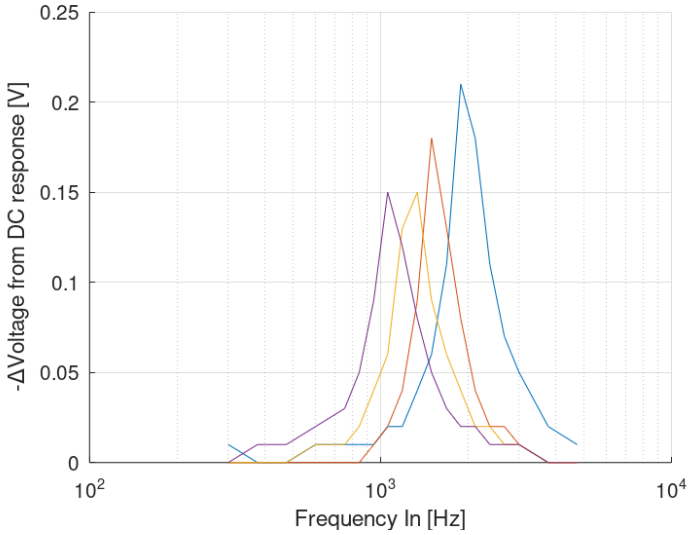

Figure 3.6: Frequency response to a V_{pp} = 0.2 V input of four C^4 + min detect CABs.

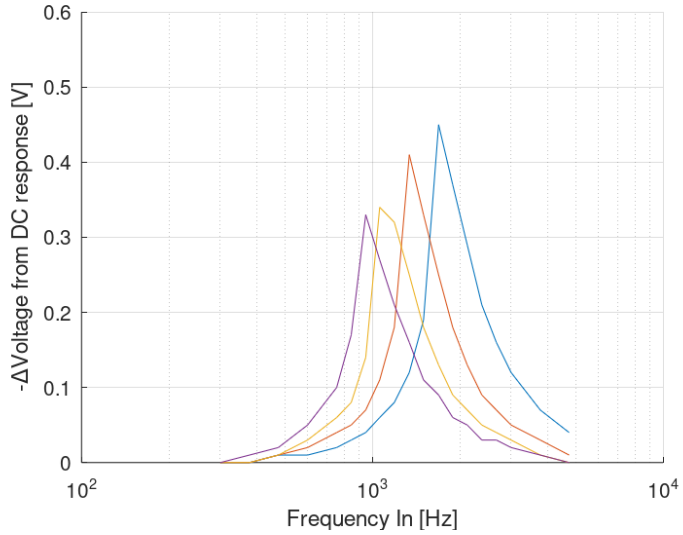

Figure 3.7: Frequency response to a V_{pp} = 0.5 V input of four C^4 + min detect CABs.

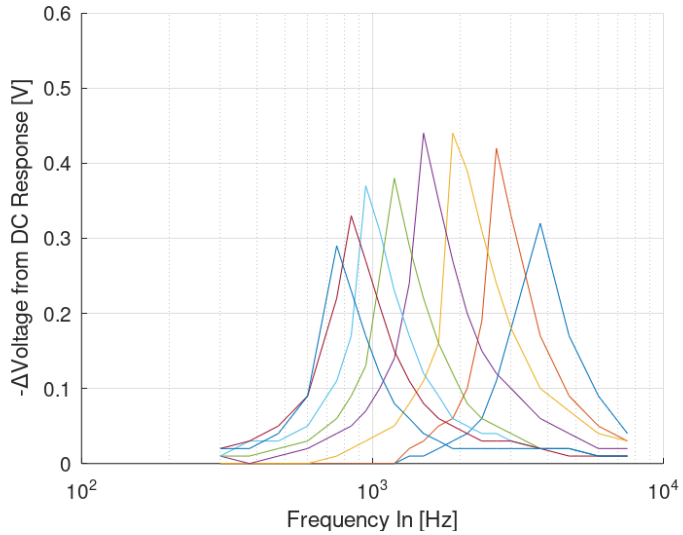

Figure 3.8: Frequency response to a V_{pp} = 0.5 V input of eight C^4 + min detect CABs.

Figure 3.8 shows the combined response of 8 C^4 + min detect CABs for locations [1;3;1] to [1;3;8] to a V_{pp} = 0.5 V input. Note that in all of these results the frequency response of the filters have all drifted upwards, but their relative spacing has not changed significantly. This is likely caused by temperature drift of the whole system.

3.2 The DSWTA + I2F CAB

3.2.1 Distributed Soft Winner-Take-All (DSWTA)

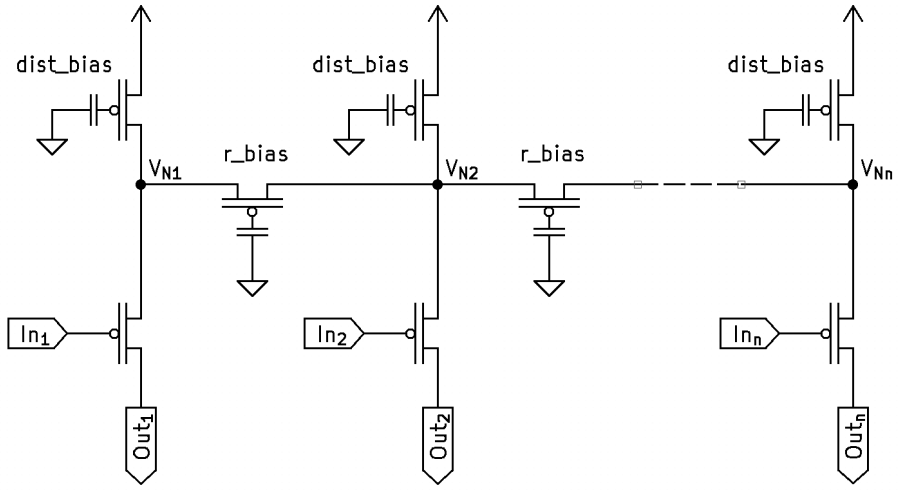

Figure 3.9: The distributed soft winner-take-all network as compiled on an FPAA.

The DSWTA circuit is responsible for computing the changing sensitivity of the human ear. Just as how a strong signal in the human ear suppresses the ability to hear weak signals nearby in frequency, a large input into the DSWTA suppresses the current output of outputs nearby in the network.

The DSWTA circuit is based off of the voltage input winner-take-all circuit, which can be seen as a generalized differential pair shown in Figure 3.10.

Let the voltage at the summing node above the biasing NFET be labeled as V_N. In that case, the current output being sinked by each NFET can be solved as

$$I_{out(m)} = I_s e^{\frac{\kappa V_{in(m)} - V_N}{U_T}},$$

$$\sum_{m=1}^{n} I_{out(m)} = I_{bias} = I_s e^{\frac{-V_N}{U_T}} \sum_{m=1}^{n} e^{\frac{\kappa V_{in(m)}}{U_T}},$$

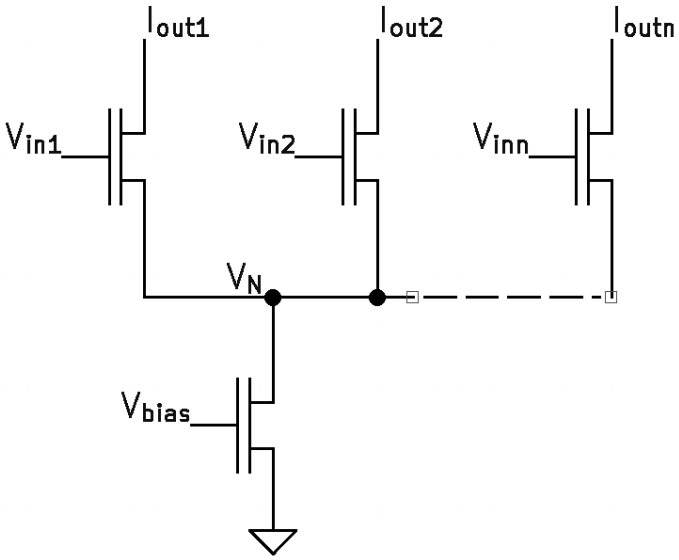

Figure 3.10: The generalized differential pair, which can be seen as an n-input, voltage input winner-take-all circuit.

$$I_{out(m)} = I_{bias} \frac{e^{\frac{\kappa V_{in(m)}}{U_T}}}{\sum_{m=1}^{n} e^{\frac{\kappa V_{in(m)}}{U_T}}}.$$

The DSWTA changes two aspects of the generalized differential pair. First, each of the input FETs is separated from each other using a diffusor element - a FG PFET in this case. The summing node is no longer a single node, but rather a series of n coupled nodes. Second, since there are now n nodes, the bias supplied to the system is distributed to each node instead of being concentrated at a single node.

For any two adjacent nodes, this new system has three different regimes of operation dependent on the values of $V_{in(m)}$, $V_{in(m+1)}$, and V_{rbias}. These regimes are caused by the use of a FG PFET as the diffusor element. Let

$$V_{CM} = (V_{in(m)} + V_{in(m+1)})/2,$$

$$\Delta V_{in(m)} = V_{in(m)} - V_{in(m+1)},$$

and V_{Nn} be the voltage at the n-th node.

The first regime of operation occurs when the FG PFET is operating in the ohmic regime. This occurs when V_{CM} is high and $\Delta V_{in(m)}$ is relatively low. In this regime, the FG PFET may be modeled as a resistor with resistance R_M dependent on $V_{in(m)}$, $V_{in(m+1)}$. For a system of two operating in this regime, shown in Figure 3.11, the current outputs can be obtained to be

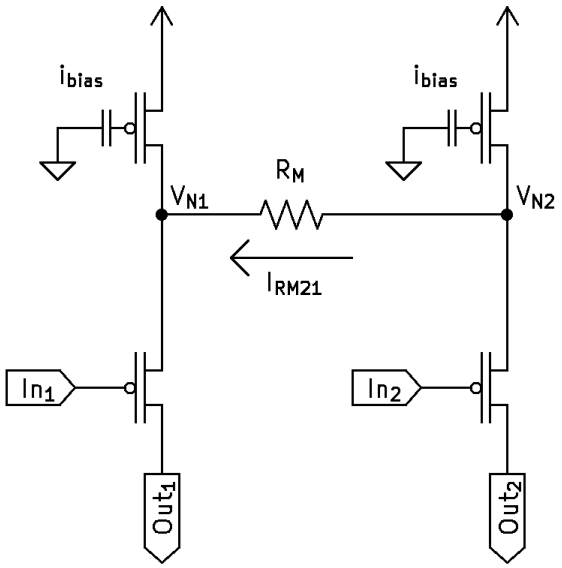

Figure 3.11: The 2-parallel distributed soft winner-take-all network operating in the ohmic regime, with the FG PFET approximated as a resistor.

$$I_m = I_s e^{\frac{-\kappa V_{in(m)} + V_{Nm}}{U_T}},$$

$$U_T \ln \frac{I_m}{I_s} = -\kappa V_{in(m)} + V_{Nm},$$

$$I_1 = I_{bias} + I_{RM21} = I_s e^{\frac{-\kappa V_1 + V_{N1}}{U_T}},$$

$$I_2 = I_{bias} - I_{RM21} = I_s e^{\frac{-\kappa V_2 + V_{N2}}{U_T}},$$

$$I_1 - I_2 = 2I_{RM21} = I_s(e^{\frac{-\kappa V_1 + V_{N1}}{U_T}} - e^{\frac{-\kappa V_2 + V_{N2}}{U_T}}) = 2\frac{V_{N2} - V_{N1}}{R_M},$$

$$V_{N2} = I_{RM21} R_M + V_{N1},$$

$$2I_{RM21} = I_s(e^{\frac{-\kappa V_1 + V_{N1}}{U_T}} - e^{\frac{-\kappa V_2 + I_{RM21}R_M + V_{N1}}{U_T}}),$$

$$I_1 + I_2 = 2I_{bias} = I_s(e^{\frac{-\kappa V_1 + V_{N1}}{U_T}} + e^{\frac{-\kappa V_2 + I_{RM21}R_M + V_{N1}}{U_T}}),$$

$$\frac{I_{RM21}}{I_{bias}} = \frac{e^{\frac{-\kappa V_1}{U_T}} - e^{\frac{-\kappa V_2 + I_{RM21}R_M}{U_T}}}{e^{\frac{-\kappa V_1}{U_T}} + e^{\frac{-\kappa V_2 + I_{RM21}R_M}{U_T}}},$$

$$\frac{I_{RM21}}{I_{bias}} = \frac{1 - e^{\frac{-\kappa(V_2 - V_1) + I_{RM21}R_M}{U_T}}}{1 + e^{\frac{-\kappa(V_2 - V_1) + I_{RM21}R_M}{U_T}}},$$

$$I_{RM21} = I_{bias} \tanh\Big(\frac{\kappa(V_2 - V_1) - I_{RM21}R_M}{2U_T}\Big).$$

Assuming operation within the linear region of $\tanh x \approx x$,

$$I_{RM21}(1 + \frac{I_{bias}R_M}{2U_T}) = I_{bias}\frac{\kappa(V_2 - V_1)}{2U_T},$$

$$I_{RM21} = I_{bias}\frac{\kappa(V_2 - V_1)}{2U_T + I_{bias}R_M},$$

$$I_1 = I_{bias}\Big(1 + \frac{\kappa(V_2 - V_1)}{2U_T + I_{bias}R_M}\Big),$$

$$I_2 = I_{bias}\Big(1 - \frac{\kappa(V_2 - V_1)}{2U_T + I_{bias}R_M}\Big).$$

The second regime of operation occurs when the FG PFET is operating in the saturation regime. This occurs when V_{CM} is high enough that the FG PFET is still conducting, and ΔV_{in} is large enough to bias the FG PFET into saturation. Let V_H be the higher of the two input voltages, and V_L be the lower of the two input voltages. The current outputs can be obtained to be

$$I_{RM} = I_{sFG}e^{\frac{-\kappa V_{rbias} + V_{NH}}{U_T}},$$

$$I_H = I_{bias} - I_{RM} = I_s e^{\frac{-\kappa V_H + V_{NH}}{U_T}},$$

$$I_{bias} = e^{\frac{V_{NH}}{U_T}}\big(I_{sFG}e^{\frac{-\kappa V_{rbias}}{U_T}} + I_s e^{\frac{-\kappa V_H}{U_T}}\big),$$

$$\ln\frac{I_{bias}}{I_s}+\frac{-V_{NH}}{U_T}=\ln\left(\frac{I_{sFG}}{I_s}e^{\frac{-\kappa V_{rbias}}{U_T}}+e^{\frac{-\kappa V_H}{U_T}}\right),$$

$$V_{NH}=-U_T\ln\left(\frac{\frac{I_{sFG}}{I_s}e^{\frac{-\kappa V_{rbias}}{U_T}}+e^{\frac{-\kappa V_H}{U_T}}}{I_{bias}/I_s}\right),$$

$$I_H=I_s\frac{e^{\frac{-\kappa V_H}{U_T}}}{\left(\frac{\frac{I_{sFG}}{I_s}e^{\frac{-\kappa V_{rbias}}{U_T}}+e^{\frac{-\kappa V_H}{U_T}}}{I_{bias}/I_s}\right)},$$

$$I_H=I_{bias}\frac{1}{1+\frac{I_{sFG}}{I_s}e^{\frac{\kappa(V_H-V_{rbias})}{U_T}}},$$

$$I_L=2I_{bias}-I_H.$$

This result is notable as it indicates that in a DSWTA network, the coupling can be limited to a maximum current dependent on the programmed coupling strength (V_{rbias}) and the neighboring input voltages.

The third regime of operation occurs when V_{CM} is low. In this regime, V_{rbias} is unable to turn on the FG PFET at all, and a negligible amount of coupling occurs: $I_m = I_{bias}$.

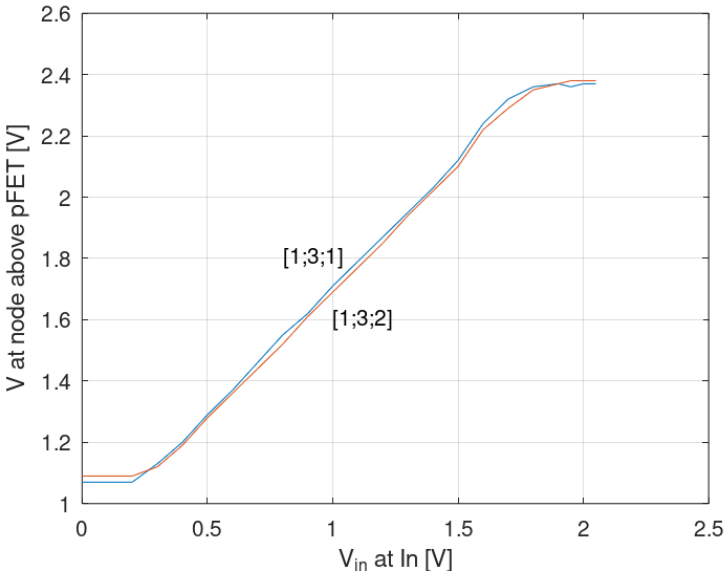

Figure 3.12: The voltage response of the DWSTA when uncoupled (r_{bias} = 0, $dist_{bias}$ = 1 nA). Note that the circuit simplifies to a PFET source follower.

Figure 3.12 shows the resulting V_{Nn} for $n = 1, 2$ when uncoupled. Note that the two

PFETs are in different CABs, but match relatively well due to their size. The slope of dV_{N1}/dV_1 and dV_{N2}/dV_2 are $\sim$0.84.

3.2.2 Current to Frequency Converter (I2F)

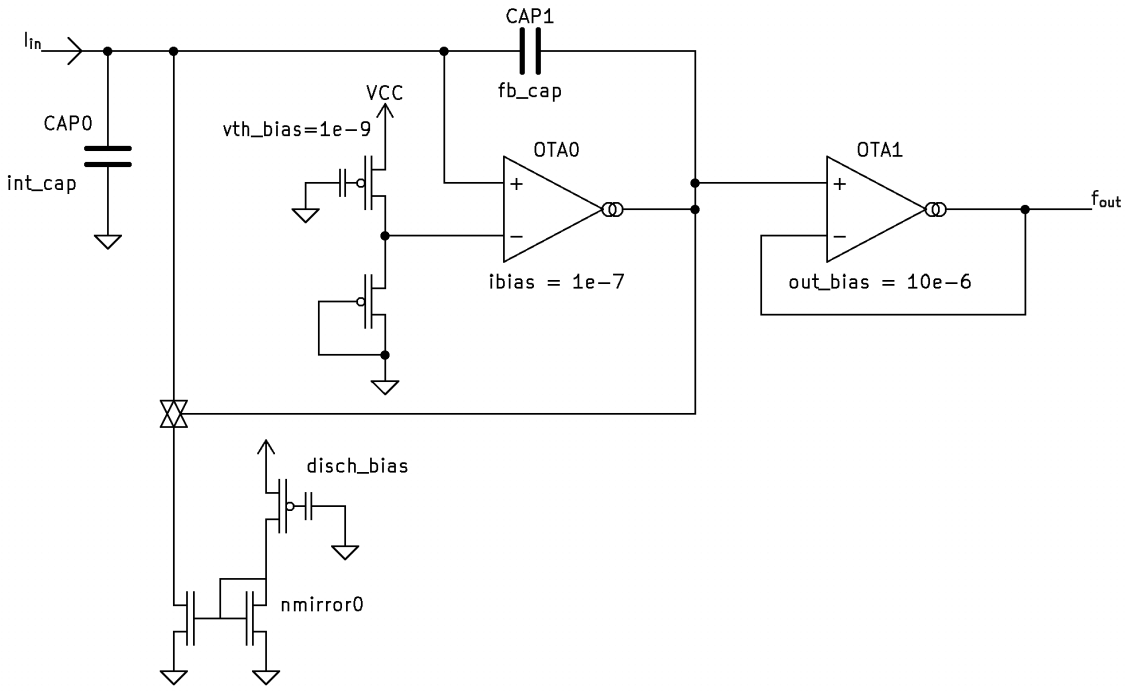

Figure 3.13: The I2F converter as compiled on an FPAA. Note that this circuit may also be viewed as an integrate-and-fire neuron, or a $\Delta\Sigma$ modulator.

The I2F block converts the analog current output of the DSWTA into a pulse density modulated (PDM) binary signal more amenable to digital signal processing. Figure 3.13 shows the schematic to the I2F converter as implemented on the FPAA, which is based off of the unclocked $\Delta\Sigma$ modulator.

This circuit works by integrating the current on the input capacitor int_cap until it reaches a threshold V_{ref}, which is set by a diode-connected PFET and a FG PFET current source. When the input voltage reaches the threshold voltage, the OTA comparator outputs high and discharges the input capacitor through a T-gate. The discharge rate is limited by an NFET mirror acting as a current source. There is also a feedback capacitor fb_cap that introduces hysteresis to the discharge process to induce oscillation.

The input to output relationship of the simple I2F converter can be described as

$$I_{in} = C_T \frac{dV_{in}}{dt},$$

$$\Delta V_{in} = V_{ref} = \frac{I_{in}}{C_T f_{out}},$$

$$f_{out} = \frac{I_{in}}{C_T V_{ref}},$$

where C_T is the total capacitance of the input node, consisting of $C_{fb} + C_{int}$. This equation assumes that when the comparator trips, C_{int} is discharged instantly to 0. Next, we assume that the discharge time is nonzero. Let T_1 be the period when the integration capacitor is charging, T_2 be the period when it is discharging, and I_{disch} be the discharge current.

$$\Delta V_{in1} = V_{ref} = \frac{I_{in}}{C_T} T_1,$$

$$\Delta V_{in2} = -V_{ref} = \frac{I_{in} - I_{disch}}{C_T} T_2,$$

$$T_2 = \frac{V_{ref} C_T}{I_{disch} - I_{in}},$$

$$\sum \Delta V_{in} = 0 = \frac{I_{in}(T_1 + T_2) - I_{disch} T_2}{C_T},$$

$$\frac{I_{in}}{f_{out}} = \frac{I_{disch} V_{ref} C_T}{I_{disch} - I_{in}},$$

$$f_{out} = \frac{I_{disch} - I_{in}}{I_{disch}} \frac{I_{in}}{C_T V_{ref}}.$$

Note that as $I_{disch} \to \infty$, this equation becomes equivalent to the instant discharge case.

The last assumption to remove is the assumption that the input voltage change rises all the way to V_{ref} and falls all the way to 0. In practice, the voltage change at the input during oscillation is caused by the feedback capacitance C_{fb} creating hysteresis through a capacitive voltage divider:

$$\Delta V_{in} = \Delta V_{out} \frac{C_{fb}}{C_{int}}.$$

Replacing V_{ref} gives

$$f_{out} = \frac{I_{disch} - I_{in}}{I_{disch}} \frac{I_{in} C_{int}}{\Delta V_{out} C_T C_{fb}},$$

$$f_{out} = \frac{I_{disch} - I_{in}}{I_{disch}} \frac{I_{in}}{\Delta V_{out} C_{fb} \left(1 + \frac{C_{fb}}{C_{int}}\right)}.$$

When $C_{fb} << C_{int}$,

$$f_{out} \approx \frac{I_{disch} - I_{in}}{I_{disch}} \frac{I_{in}}{\Delta V_{out} C_{fb}}.$$

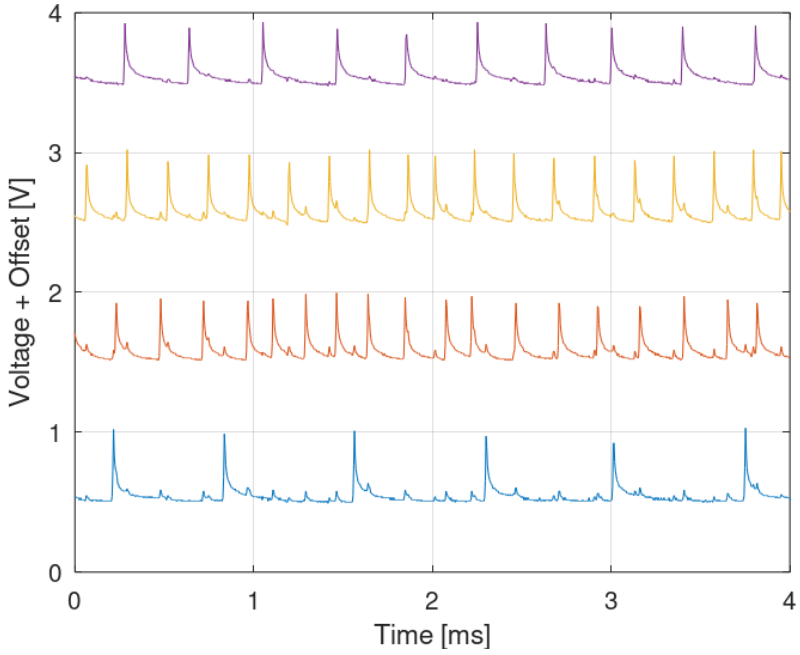

Figure 3.14: Typical output waveforms for four programmed I2F converters. The plots have been offset for clarity. Note the presence of small spikes in each output due to parasitic capacitive coupling.

When implemented on the FPAA, there is a large capacitance at the drain of the NFET mirror that causes a current much larger than I_{disch} to flow through the T-gate during the discharge phase. This causes the I2F converter to behave closer to the simplified $\Delta\Sigma$ modulator with instant discharge. In this regime of operation, ΔV_{out} is also observed to be relatively constant at $\sim$0.5 V, shown in Figure 3.14. Thus, the frequency output of the I2F converter is demonstrated to be a relatively linear function of the input current.

The output is buffered through another OTA to improve noise rejection feeding in from

f_{out}. Figure 3.14 shows capacitive coupling of the outputs of four I2F converters. Without buffering, this capacitive coupling can cause premature spiking by feeding through C_{fb} and causing V_{in} to rise above V_{ref}. This causes synchronization of two or more I2F converters if their frequencies are similar.

The desired output frequency of the I2F converter is determined by the output precision desired for each frame used in encoding. For the MP3 frame length of ∼26 ms, obtaining 8 bits of precision for each frame requires a spiking frequency range of $2^8/T \approx 9.8$ kHz. In implementation, the I2F converter block is observed to be able to easily reach frequencies above ∼80 kHz, with the primary limitation being the drive capability of the OTA buffer and various parasitic capacitances causing attenuation of the output signal.

3.2.3 CAB Behavior

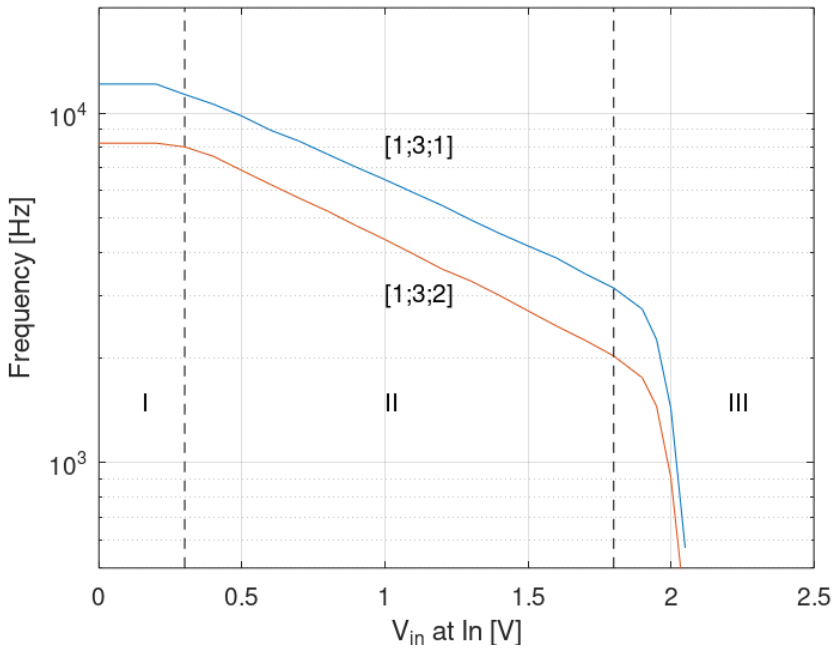

Figure 3.15: Response of the uncoupled DSWTA + I2F CAB compiled at two different locations. Note that there is a large log-linear region in voltage to frequency conversion from 0.3 V to 1.8 V. No cross-dependence of f_{out1} on V_{in2} was observed, and vice versa.

Figure 3.15 shows the frequency output response of the I2F converter as a function

of input voltages to an uncoupled DSWTA. Without any coupling, the output of the I2F converter is directly dependent on the current going through the FG PFET. It shows three clear regions of operation with breakpoints at V_{in} ∼0.3 V and V_{in} ∼1.8 V.

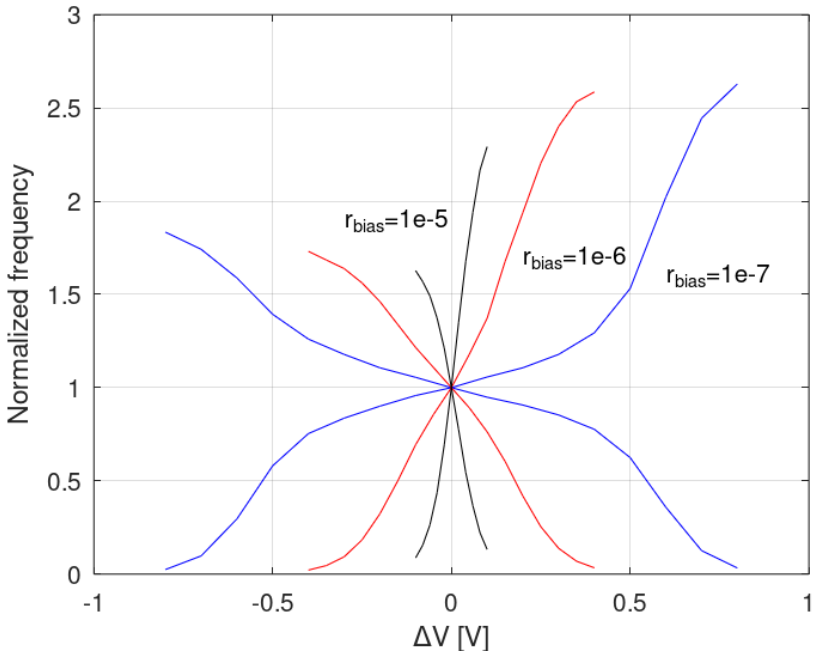

Figure 3.16: Response of two cross-coupled DSWTA + I2F CABs for three different values of r_{bias} at V_{CM} = 1.5 V. Differences in V_{in} cause a differential shift in f_{out}, which has been normalized to $\Delta V = 0$.

In region I where V_{in} <∼0.3 V, the observed frequency output of the I2F converter is constant. In this region, V_{in} is low enough to cause $V_N < V_{ref}$, and a maximum frequency limit of the I2F converter is observed when only one FG PFET current source feeds into the I2F converter.

In region II where ∼0.3 V< V_{in} <∼1.8 V, the frequency output becomes a log-linear function of the voltage input. This is caused by the parasitic capacitance between V_N and the gate of the FG PFET being used as the current source. Due to this parasitic capacitance, there is a capacitive voltage divider causing a small $\Delta V_{FGgn}/\Delta V_{Nn}$. However, since the FG PFET is operating in subthreshold, the current I_{ds} increases exponentially from this small ΔV_{FGgn}, causing the log-linear response.

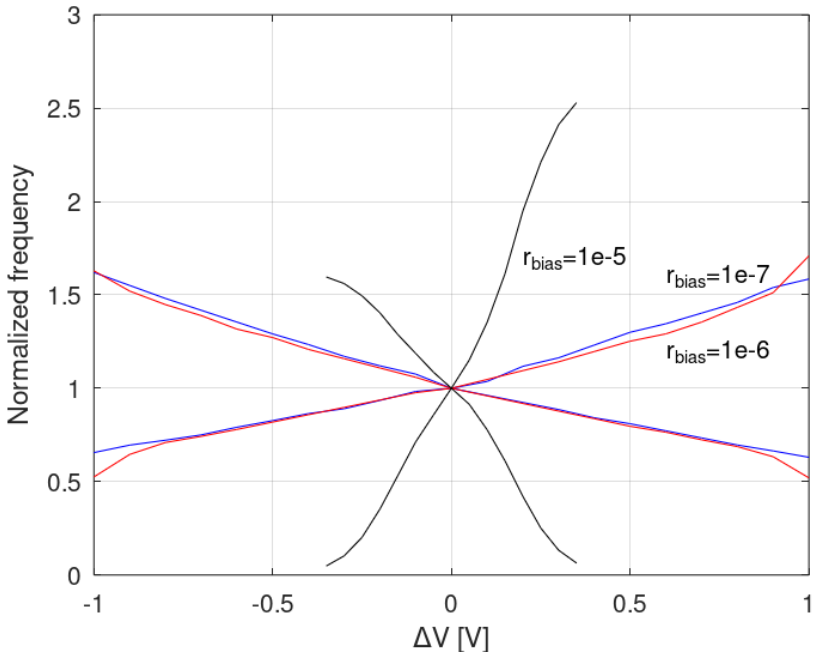

Figure 3.17: Response of two cross-coupled DSWTA + I2F CABs for three different values of r_{bias} at V_{CM} = 1 V. Differences in V_{in} cause a differential shift in f_{out}, which has been normalized to $\Delta V = 0$.

In region III where V_{in} >∼1.8 V, V_N is observed to approach V_{DD} =2.5 V from Figure 3.12. In this region, the FG PFET is no longer operating in saturation mode, causing a significant drop in supplied current as V_{in} increases.

The desired operation regime is in region II, where the I2F converter has large and predictable log-linear response. It also allows for an estimate of V_{in}, which is useful for estimating an overall strength of the input signal. In other words, this creates a dependence on the V_{CM} of the DSWTA inputs that gives an indication of the strength across all inputs, even if all inputs are equal. If region II was flat like region I, then this dependence on V_{CM} is destroyed, and the ability to distinguish between having a large equivalent voltage on all nodes versus having a small equivalent voltage on all nodes is lost. Additionally, it allows for an estimation of V_{in} if coupling is lost due to low V_{CM} on the DSWTA inputs.

Note that there is a mismatch in the natural oscillation frequency of the two I2F converters. This is primarily caused by mismatch in C_{fb} between the different CABs, as C_{fb} is

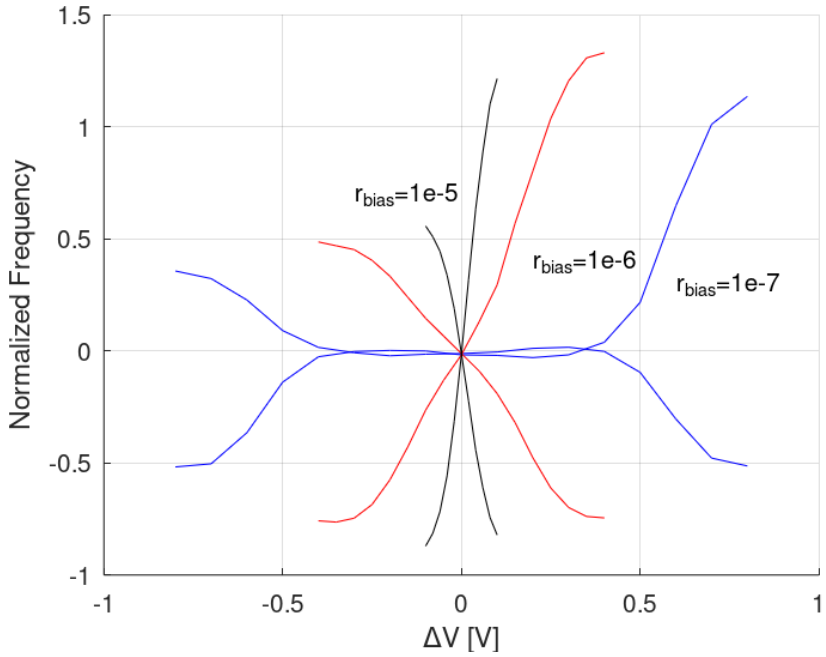

Figure 3.18: Response of two cross-coupled DSWTA + I2F CABs for three different values of r_{bias} at V_{CM} = 1.5 V. The intrinsic $f_{out}(V_{in})$ slope has been fitted and removed to better observe the coupling behavior. Note the absence of coupling between ∼–0.4 V to ∼0.4 V in the r_{bias} = 0.1 µA case.

primarily determined by parasitic capacitance inside the CAB routing. To adjust for this inherent mismatch, the frequency outputs of the I2F converters in Figure 3.16 to Figure 3.19 are divided by their natural oscillation frequencies at $\Delta V_{in} = 0$.

Figure 3.16 shows the effect of differing programmed values of r_{bias} at V_{CM} =1.5 V. Note that r_{bias} indicates the programmed current of the FG PFET used as the diffusor element, and that the asymmetry in maximum frequencies achieved by the I2F converters is caused by the mismatch of C_{fb}.

Recall that the frequency outputs of the I2F converters are dependent on V_{in} even without coupling. To better observe the cross-coupling effect, this intrinsic dependence was removed by taking single-ended data of the two I2F converters without any coupling, and then subtracting the dependence from the results in Figure 3.16 and Figure 3.17 to obtain Figure 3.18 and Figure 3.19, respectively.

In Figure 3.18, removing the intrinsic $f_{out}(V_{in})$ slope shows a distinct lack of coupling

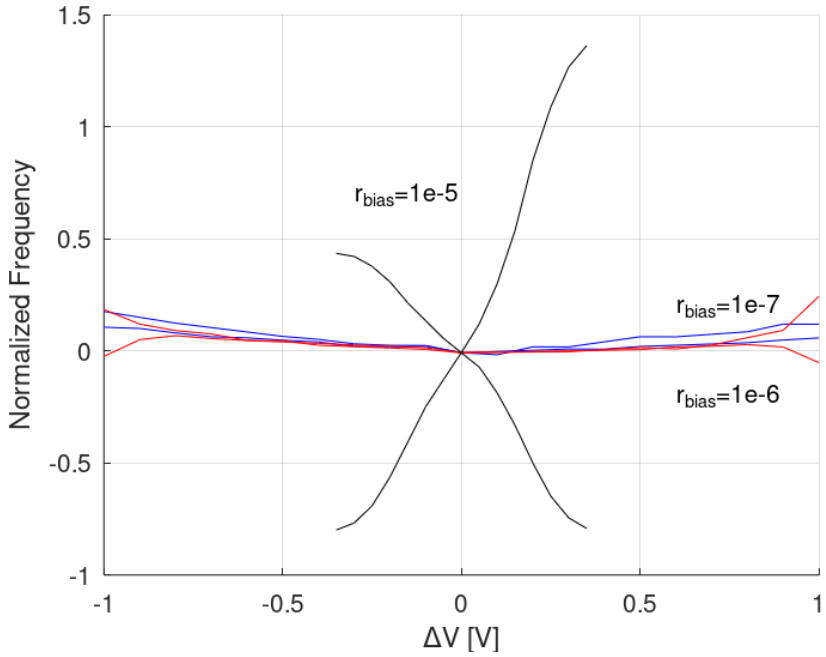

Figure 3.19: Response of two cross-coupled DSWTA + I2F CABs for three different values of r_{bias} at V_{CM} = 1 V. The intrinsic $f_{out}(V_{in})$ slope has been fitted and removed to better observe the coupling behavior. Note the absence of coupling in the r_{bias} = 0.1 µA case, and the absence of coupling between ∼–0.9 V to ∼0.9 V in the $r_{bias} = 1$ µA case.

behavior between ∼–0.4 V to ∼0.4 V in the r_{bias} = 0.1 µA case. The voltage programmed onto the coupling FG PFET in the DSWTA is not low enough to cause a measurable amount of current to flow between V_{N1} and V_{N2}. As $|\Delta V|$ increases above ∼0.4 V, $\max(V_{N1}, V_{N2})$ rises high enough to allow current to conduct between V_{N1} and V_{N2}.

Similarly, in Figure 3.19, removing the intrinsic $f_{out}(V_{in})$ slope shows a complete lack of coupling behavior in the r_{bias} = 0.1 µA case across all tested ΔV_{in}. For r_{bias} = 1 µA, a small but detectable amount of coupling behavior can be observed near $|\Delta V|$ = 1 V.

In both Figure 3.18 and Figure 3.19, coupling behavior is observed in the r_{bias} = 10 µA case with a distinct decrease in coupling strength as V_{CM} is changed from 1.5 V to 1 V. This result agrees with the analysis of the DSWTA in subsection 3.2.1.

3.3 Processing the PDM Signal

3.3.1 Preamplifier and Counter

To process the observed PDM signal into a binary value, the signal must first be amplified to an appropriate level. Figure 3.20 shows the proposed preamplifier structure consisting of a comparator and a lowpass OTA-C filter with a tunable time constant. An OTA-C filter is used instead of a fixed reference voltage as the DC level of the I2F converter is observed to shift as a function of output frequency. The OTA-C filter filters out the unwanted low frequency signal. The comparator is used to obtain a quick rail-to-rail output and high gain for input into digital blocks.

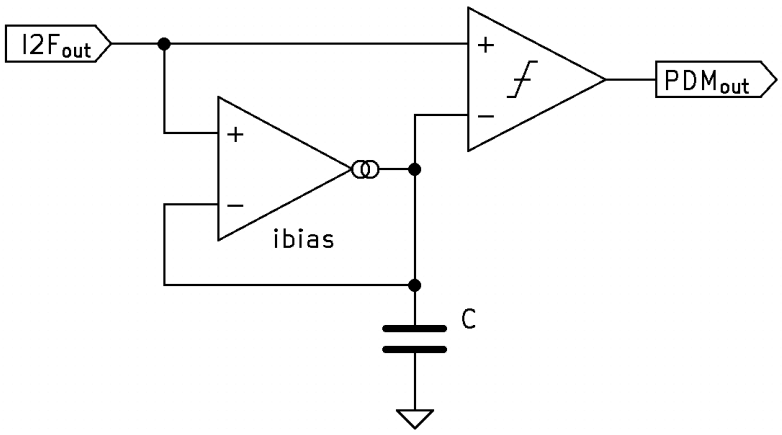

Figure 3.20: Proposed preamplifier block for converting the raw I2F output into a pulse density modulated digital signal.

To turn the PDM signal into a binary value, the signal is fed into the clock of an n-bit counter. The number of bits of the counter is determined by the maximum output frequency of the I2F converter, which is in turn determined by the desired output precision per frame. The minimum required bit counter required can be calculated using $n + \log_2 P$, where n is desired the output bit precision and P is the parallelization number. In practice, the output frequency of the I2F converter is greater than the frequency needed to obtain n bits of desired output precision. Therefore, the minimum required bit counter is calculated to be

$\log_2 f_{max}T_{frame} + \log_2 P$. Referencing Figure 3.15 to obtain a conservative f_{max} = 20 kHz, the 8-parallel system would require at least a 12-bit counter on each output for a given frame length of 26 ms.

3.3.2 Exponential to Linear Spacing Converter

This section seeks to address the interface between the exponentially spaced filterbank used in the analog psychoacoustic model and the linearly spaced MDCT. For M bands exponentially spaced from f_{min} to f_{max}, the center frequency of the m-th filter can be calculated as

$$f_{c,m} = f_{min}\left(\frac{f_{max}}{f_{min}}\right)^{\frac{m-0.5}{M}}.$$

Appropriately tuned, the frequency response of adjacent filters will intercept each other at

$$f_{x(m,m+1)} = f_{min}\left(\frac{f_{max}}{f_{min}}\right)^{\frac{m}{M}}.$$

Let $H_m(s)$ be the transfer function of the m-th filter. Let $N_m(s)$ be the normalized transfer function of the m-th filter as defined by

$$N_m(s) = \frac{H_m(s)}{\int_{s_{min}}^{s_{max}} |H_m(s')|ds'},$$

where $s_{min} = j2\pi f_{min}$ and $s_{max} = j2\pi f_{max}$. The amount that the m-th exponentially spaced band falls within the n-th linearly spaced band is

$$C_{mn} = \left| \int_{s_{lin,n-}}^{s_{lin,n+}} N_m(s'')ds'' \right|,$$

where $s_{lin,n-} = j2\pi f_{lin,n-}$ and $s_{lin,n+} = j2\pi f_{lin,n+}$, representing the lower and upper frequencies of the n-th linearly spaced band:

$$f_{lin,n-} = f_{min} + (f_{max} - f_{min})\frac{n-1}{M},$$

$$f_{lin,n+} = f_{min} + (f_{max} - f_{min})\frac{n}{M}.$$

Therefore, for each linearly spaced band in the DCT, the contribution of each exponential band can be weighted by a precalculated constant determined by the amount the exponential band falls within the linearly spaced band.

For the built 8-parallel filterbank with parameters f_{min} = 300 Hz, f_{max} = 3 kHz shown in Figure 3.3, after adjusting for temperature drift, the following C_{mn} values were calculated and shown in Table 3.2.

Table 3.2: Calculated Exponential to Linear Converter Constants C_{mn} for Built 8-Parallel System

	n=1	n=2	n=3	n=4	n=5	n=6	n=7	n=8
m=1	0.48856	0.14982	0.09419	0.07232	0.05337	0.05294	0.05333	0.03546
m=2	0.45826	0.17546	0.09984	0.07385	0.05367	0.05223	0.05216	0.03454
m=3	0.25925	0.30752	0.13129	0.08808	0.06091	0.05854	0.05708	0.03734
m=4	0.11690	0.33646	0.19166	0.11025	0.07204	0.06725	0.06392	0.04154
m=5	0.06174	0.17017	0.30437	0.16450	0.09456	0.08198	0.07504	0.04765
m=6	0.03471	0.07053	0.15910	0.28354	0.16487	0.12330	0.10268	0.06127
m=7	0.02133	0.03864	0.06584	0.12526	0.22386	0.25124	0.17902	0.09482
m=8	0.01641	0.02807	0.04396	0.06894	0.09797	0.19712	0.34554	0.20199

Note that in the built system, the CAB locations are inverted relative to the m-th exponential converter; the highest frequency C^4 bandpass filter corresponding to m = 8 is compiled and programmed at location [1;1;1], and the lowest frequency C^4 bandpass filter corresponding to m = 1 is compiled and programmed at location [1;1;8].

3.4 Full System Test Results

The full system from analog audio input to the output of the I2F converters was built on the FPAA and characterized for select values of V_{CM}, r_{bias}, and $V_{in,pp}$. For each system, the outputs of the min detects were first tuned to output 1.5 V when V_{in}=1.25 VDC and V_{bias} of the bandpass filter block = 1.5 V. Note that V_{bias} of the bandpass filter blocks are connected to V_{CM}. For the 2-parallel and 4-parallel system, dist_bias = 1 nA and disch_bias = 4 nA. For the 8-parallel system, dist_bias = 2 nA and disch_bias = 8 nA, causing an approximate doubling of the I2F output frequencies.

3.4.1 The 2-Parallel System

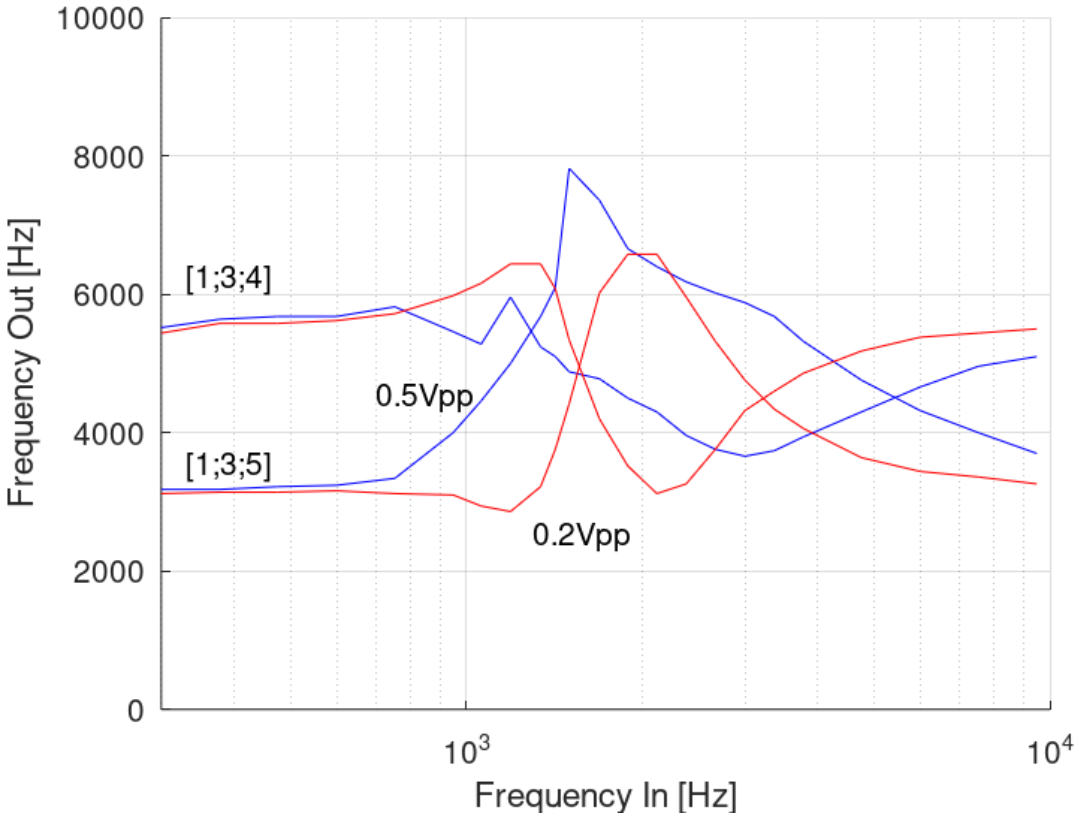

Figure 3.21: Frequency in, frequency out plot of the 2-parallel full system at V_{CM} = 1.25 V, r_{bias} = 7 µA. Each color corresponds to a different input amplitude. Note that the coupling strength is a function of both input amplitude and DC level.

Figure 3.21 shows the frequency in, frequency out plot of the 2-parallel system at V_{CM} = 1.25 V, r_{bias} = 7 µA for two different input voltage amplitudes. Note that the coupling behavior is reduced in the blue V_{pp} = 0.5 V traces, as the larger V_{pp} causes a lower V_{CM}. The two sharp peaks in the blue traces occur at the peak gain responses of the C^4 bandpass filters.

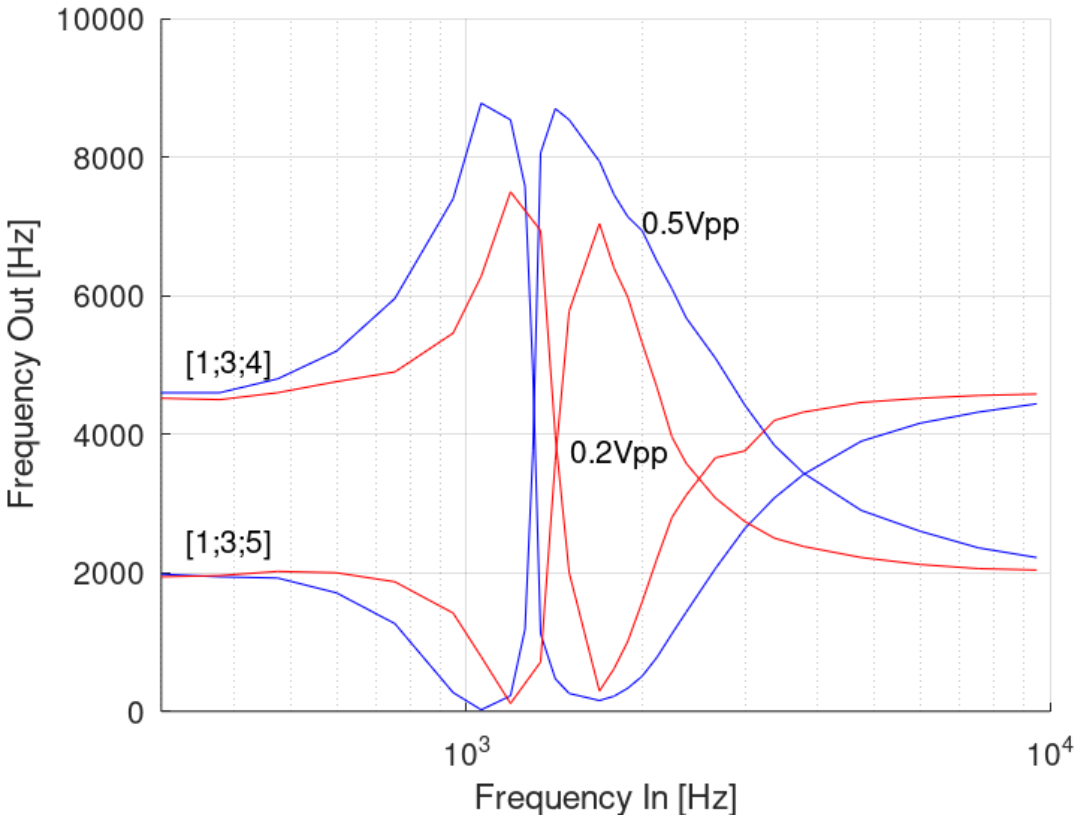

Figure 3.22: Frequency in, frequency out plot of the 2-parallel full system at V_{CM} = 1.5 V, r_{bias} = 7 µA. Each color corresponds to a different input amplitude. Note that the coupling strength is a function of both input amplitude and DC level.

In the red V_{pp} = 0.2 V traces, the coupling is stronger because of the increased V_{CM}. The frequency outputs of the I2F converters show a clear coupling effect: as frequency increases, the [1;3;4] I2F converter increases and a small suppressive effect is seen on the [1;3;5] I2F converter. As the frequency increases above ∼1.2 kHz, the DSWTA effect kicks in as the [1;3;5] begins to "win", suppressing the [1;3;4] I2F converter. As the frequency increases further above ∼2 kHz, the system slowly returns to the DC response as the input

frequency goes into the stopband of both C^4 filters.

Figure 3.22 shows the frequency in, frequency out plot of the 2-parallel system at V_{CM} = 1.5 V, r_{bias} = 7 µA for two different input voltage amplitudes. At this bias point, the higher V_{CM} causes a much stronger coupling to occur in the DSWTA. This effect is seen in the I2F outputs where a small change in frequency causes a significant change in the frequency outputs, especially between the corner frequencies of the two C^4 filters. The stronger coupling also causes a much stronger suppressive effect as the "winning" input into the DSWTA steals more current from the "losing" input.

Additionally, the greater V_{in} causes an increase in coupling strength. This is in contrast to the V_{CM} = 1.25 V case, where due to the low V_{CM}, a decrease in coupling strength was seen as the inputs to the DSWTA became too low and coupling strength was lost.

3.4.2 The 4-Parallel System

Figure 3.23 shows the frequency in, frequency out plot of the 4-parallel system at V_{CM} = 1.5 V, V_{in} = 0.1 V, r_{bias} = 7 µA, compiled for I2F CABs [1;3;4] to [1;3;7] fed from C^4 + min detect CABs [1;1;4] to [1;1;7]. As f_{in} increases, each of the C^4 + min detect CABs takes its turn "winning" in the DSWTA and suppressing the other CABs.

Figure 3.24 shows the same response when V_{in} is increased to 0.2 V, causing more significant "winning" in the DSWTA. This is particularly evident near f_{in} = 2 kHz where the large response of the [1;1;4] min detect, coupled with the strong coupling, significantly suppresses all other I2F converters. Here, the output of the [1;3;7] I2F converter was suppressed below measuring capabilities; all spiking seen coming from [1;3;7] in this case was triggered by parasitic capacitance coupling from other outputs, and no regular spiking was detectable.

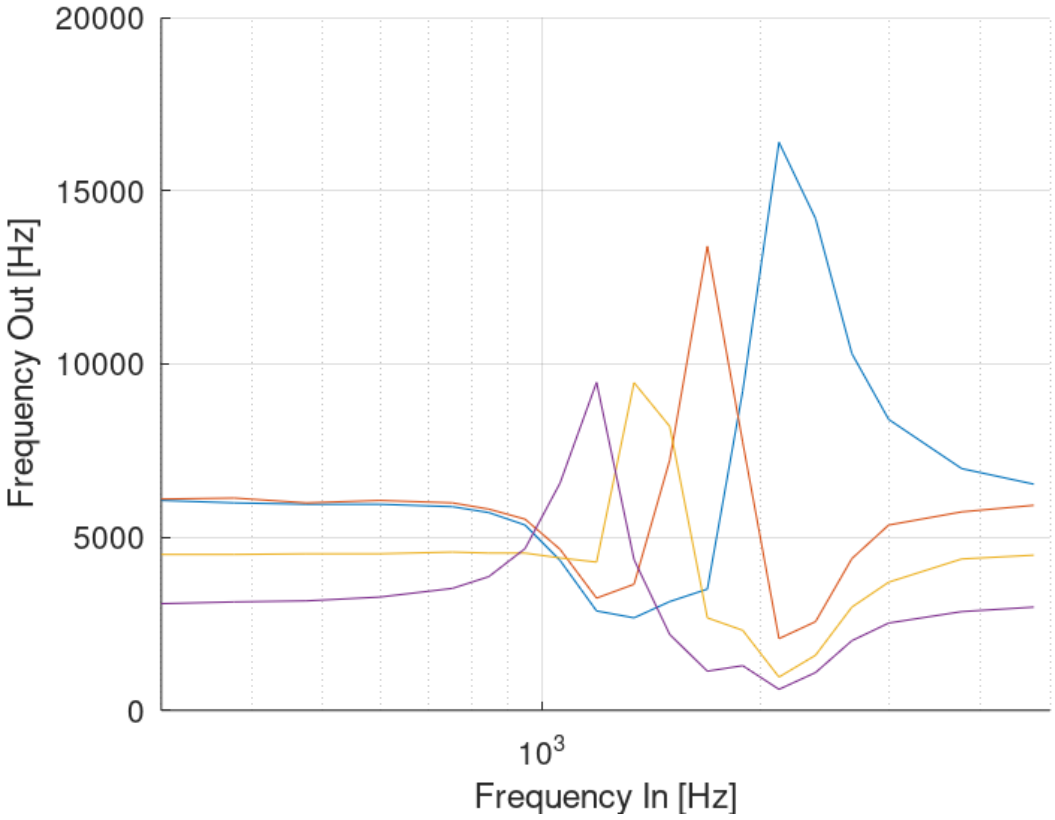

Figure 3.23: Frequency in, frequency out plot of the 4-parallel full system at V_{CM} = 1.5 V, V_{in} = 0.1 V, r_{bias} = 7 µA.

3.4.3 The 8-Parallel System

Figure 3.25 shows the frequency in, frequency out plot of the 8-parallel system at V_{CM} = 1.5 V, V_{in} = 0.5 V, r_{bias} = 4 µA. Based off of the results of the 4-parallel system, a lower r_{bias} was chosen to reduce the coupling, "softening" the DSWTA. Similar to the 4-parallel system, as f_{in} increases, each CAB takes its turn "winning" the DSWTA and suppressing the other CABs, with CABs nearby suppressed to a greater extent than CABs further away.

One interesting effect is the jaggedness observed in most of the I2F CAB outputs. This is caused by a combination of the weak r_{bias} = 4 µA coupling and the large V_{in}. Due to the 2nd-order nature and relatively low $Q \approx 3$ of the C^4 bandpass filters, the stopband roll-off of the filterbank is not extremely sharp. The C^4 bandpass filter is still responding significantly from the input frequency even if it is a band or two away in center frequency. This

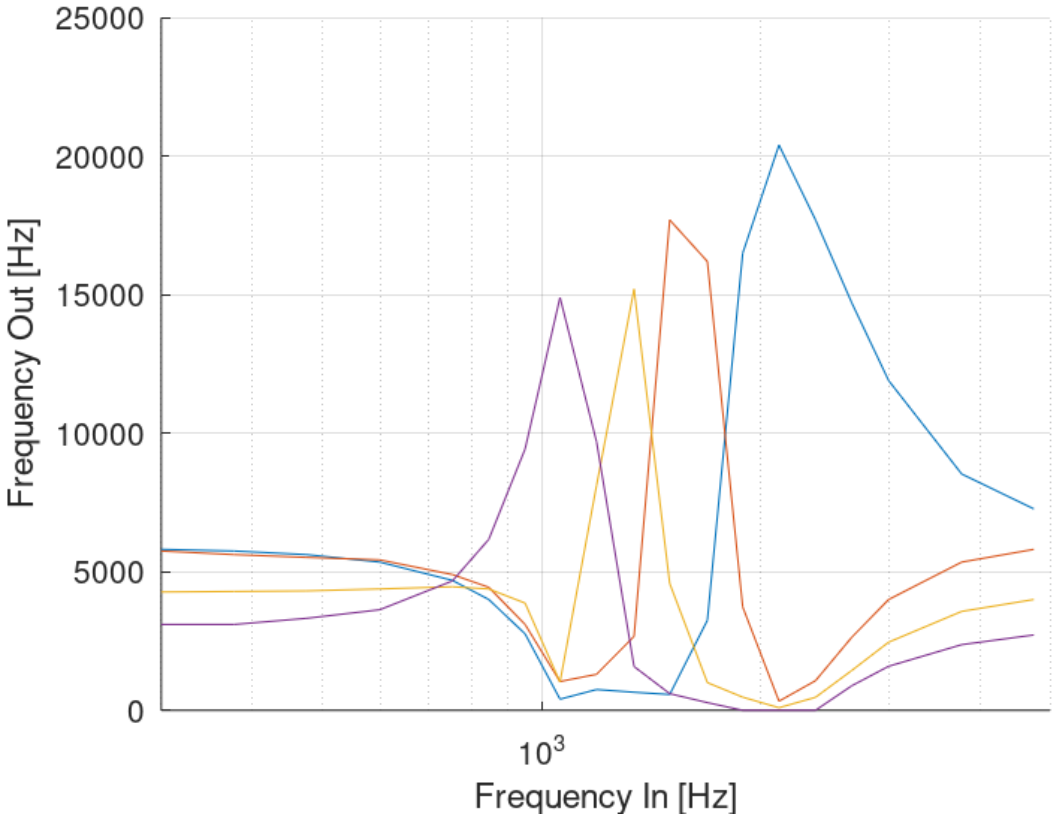

Figure 3.24: Frequency in, frequency out plot of the 4-parallel full system at V_{CM} = 1.5 V, V_{in} = 0.2 V, r_{bias} = 7 µA. Note the stronger suppressive effect due to the stronger winner-take-all effect compared to the V_{in} = 0.1 V case.

combines with the coupling strength decrease phenomenon when two adjacent DSWTA inputs are both low to cause a splitting effect in the DSWTA, hence the jaggedness in the frequency output graph. This effect should not occur at a higher r_{bias} or lower V_{in} and is a result of the unique biasing circumstances.

Figure 3.26 shows the response of the system to a f_{in} = 754 Hz sine, triangle, and square wave input. The [1;3;6] I2F converter is centered at 754 Hz and shows a clear response to all input signals. Nearby CABs are suppressed, and the aforementioned DSWTA splitting behavior is also seen. The [1;3;2] I2F converter is strongly excited by the 3rd harmonic component of the square wave signal and causes two winners in the DSWTA response. Similarly, the presence of the 3rd harmonic component in the triangle wave response causes a small but clearly discernible difference compared to the sine wave response.

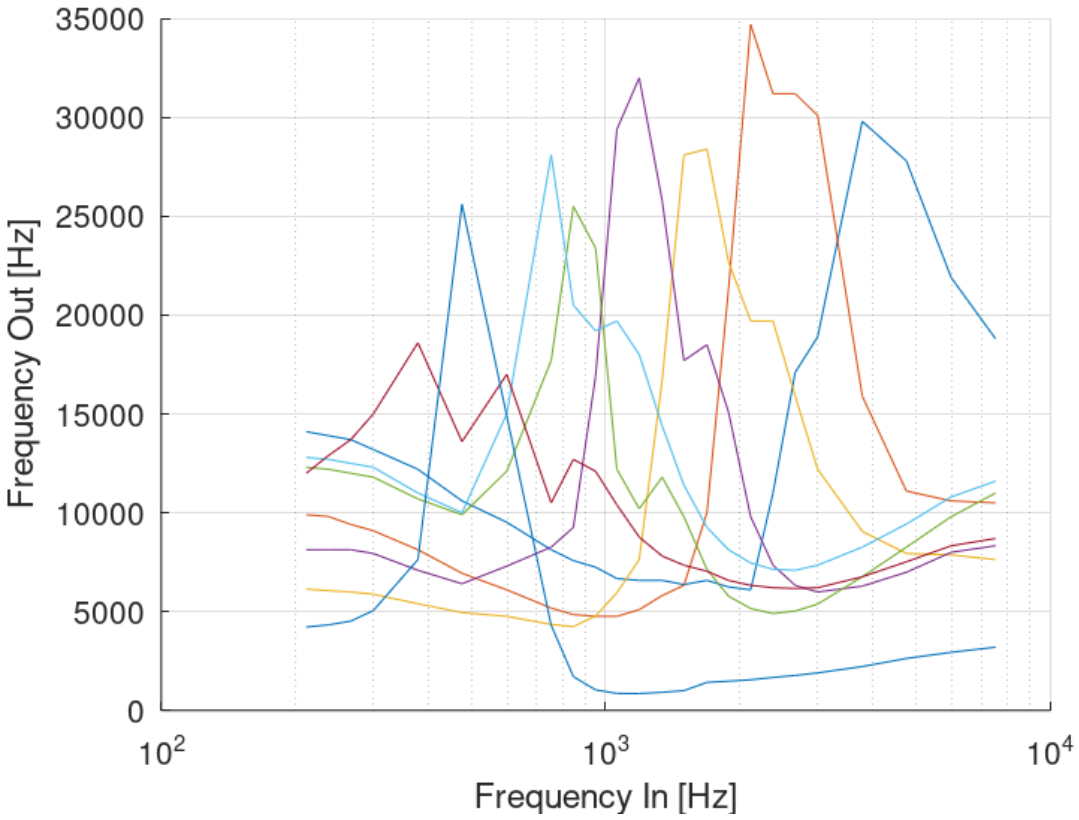

Figure 3.25: Frequency in, frequency out plot of the 8-parallel full system at V_{CM} = 1.5 V, V_{in} = 0.5 V, r_{bias} = 4 µA.

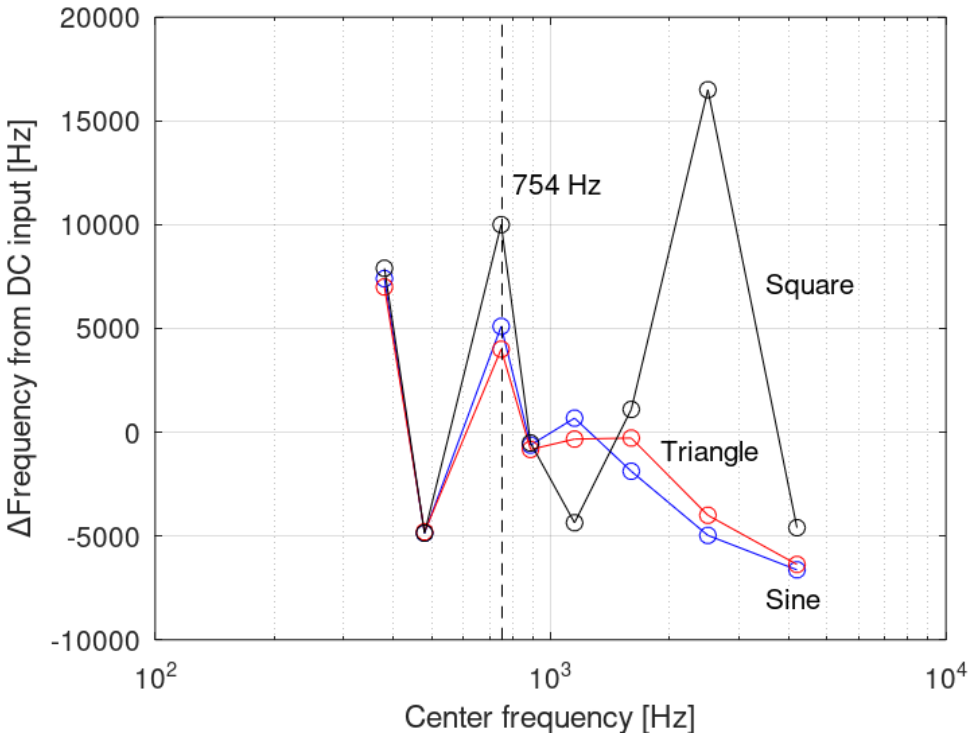

Figure 3.26: Frequency out plot of the 8-parallel full system at V_{CM} = 1.5 V, V_{in} = 0.5 V, r_{bias} = 4 µA, f_{in} = 754 Hz for a sine, triangle, and square wave inputs. Note the two prominent peaks of the square wave input corresponding to its fundamental and 3rd harmonic components, and the slight deviation of the triangle response to the sine response.

CHAPTER 4
INTEGRATION

4.1 Summary

Chapter 1 explained the psychoacoustic phenomenon of auditory masking that allows for lossy audio compression to occur without a perceptible loss in audio quality. Chapter 2 explained the device used for implementing the analog psychoacoustic model as well as the process for programming the system onto the device. Chapter 3 analyzed each of the building blocks in the system and showed experimental results for each of the blocks. Chapter 3 also showed experimental results and analysis of cascaded building blocks and of the full system. Additionally, Chapter 3 provided a possible converter from the exponentially spaced output of the system to the linearly spaced MDCT. The primary work done centered around the design, creation, and characterization of each building block in the analog psychoacoustic model as well as the creation and characterization of the full system.

4.2 Power Consumption Estimate

The total power consumption of the analog psychoacoustic model can be roughly estimated by summing all of the bias currents present in all blocks and multiplying by the power supply voltage of 2.5 V. Table 4.1 tallies the current consumption of each system block for the implemented 8-parallel full system.

Table 4.1: Current consumption estimate of each block in the implemented 8-parallel system

System Block	Current Consumption
C^4 Filter Bank	39.4 nA
Min Detect	80.8 nA
DSWTA	16 nA
I2F	872 nA
Output Buffers	160 µA
Total	161 µA
Total (without output buffers)	1.01 µA

Note that the total current consumption is dominated by the buffers used to buffer signals coming out of each CAB. The current consumption of these buffers may be reduced or removed if the signal were to be processed inside the chip, giving a total power consumption of $\sim$2.5 µW for the computational elements of the full 8-parallel system. The power consumption scales roughly linearly with the parallelization number as it is dominated by the I2F converter block. Using this approximation gives an estimated power consumption of 10 µW for a theoretical 32-parallel system that spans from 20 Hz to 20 kHz if implemented on an FPAA.

REFERENCES

[1] ISO/TC 43 Acoustics, "Acoustics – Normal equal-loudness-level contours," International Organization for Standardization, Geneva, CH, Standard ISO 226:2003, 2003.

[2] W. M. Leach, "Basic principles of sound," in *Introduction to Electroacoustics and Audio Amplifier Design*. Kendall Hunt Publishing Company, 2003, pp. 1–12.

[3] I. Peretz and K. L. Hyde, "What is specific to music processing? Insights from congenital amusia," *Trends in Cognitive Sciences*, vol. 7, no. 8, pp. 362–367, Aug. 2003.

[4] M. B. C. J., *Cochlear hearing loss: Physiological, psychological and technical issues*. Wiley Blackwell, 2008.

[5] M. Bazyar and R. Sudirman, *A Robust Data Embedding Method for MPEG Layer III Audio Steganography*, Dec. 2015.

[6] S. Suh, A. Basu, C. Schlottmann, P. E. Hasler, and J. R. Barry, "Low-power discrete fourier transform for ofdm: A programmable analog approach," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 58, no. 2, pp. 290–298, 2011.

[7] D. García Moreno, A. A. Del Barrio, G. Botella, and J. Hasler, "A cluster of fpaas to recognize images using neural networks," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 68, no. 11, pp. 3391–3395, 2021.

[8] H. Jeong, J. Kim, and W.-k. Cho, "Low-power multiplierless dct architecture using image correlation," *IEEE Transactions on Consumer Electronics*, vol. 50, no. 1, pp. 262–267, 2004.

[9] *MP3 / OGG Vorbis Encoder and Audio Codec Circuit*, VS1063a, Version 1.31, VLSI Solution, Oct. 2017.

[10] P. Hasler, B. Minch, and C. Diorio, "Adaptive circuits using pFET floating-gate devices," in *Proceedings 20th Anniversary Conference on Advanced Research in VLSI*, 1999, pp. 215–229.

[11] J. Hasler, *SoC FPAA Routing Discussion*, 2016.

[12] J. O. Hasler and A. Natarajan, "An Open-Source ToolSet for FPAA Design," 2020.

[13] D. W. Graham, P. E. Hasler, R. Chawla, and P. D. Smith, "A low-power programmable bandpass filter section for higher order filter applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 6, pp. 1165–1176, 2007.

[14] S. Shah, H. Toreyin, J. Hasler, and A. Natarajan, "Temperature sensitivity and compensation on a reconfigurable platform," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 26, no. 3, pp. 604–607, 2018.

[15] S. Shah and J. Hasler, "Low power speech detector on a FPAA," in *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2017, pp. 1–4.