ANALOG ADAPTIVE NONLINEAR FILTERING AND SPECTRAL ANALYSIS FOR LOW-POWER AUDIO APPLICATIONS

by

Christopher D. Salthouse

S.B. Electrical Engineering

Massachusetts Institute of Technology, 2000
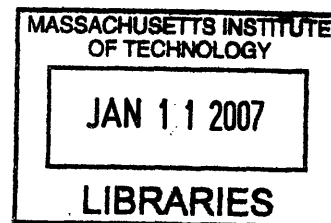
M.Eng Electrical Engineering

Massachusetts Institute of Technology, 2000

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN ELECTRICAL ENGINEERING AT THE

MASSAHCUSETTS INSTITUTE OF TECHNOLOGY
[September 2006]
AUGUST 2006

Signature of the Author: _____

Department of Electrical Engineering

August 8th, 2006

Certified by: _____

Rahul Sarpeshkar

Associate Professor of Electrical Engineering

Thesis Supervisor

Accepted by: _____

Arthur C. Smith

Chairman, Department Committee on Graduate These

ANALOG ADAPTIVE NONLINEAR FILTERING AND SPECTRAL ANALYSIS FOR LOW-POWER AUDIO APPLICATIONS

by

Christopher D. Salthouse

Submitted to the Department of Electrical Engineering and Computer Science on
August 8th, 2006 in Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy in Electrical Engineering

Abstract

Filters are one of the basic building blocks of analog circuits. For linear operation, the power consumption is proportional to the dynamic range for a given topology. I have explored techniques to lower the power consumption below this limit by extending operation beyond the linear range.

First, I built a power-efficient linear gm-C filter that demonstrates that dynamic range can be shifted to higher linear ranges using capacitive attenuation. In a standard gm-C filter, the minimum noise is limited by the discrete charge on the electrons and holes stored on the capacitor. This noise can only be reduced by collecting more charge on a larger capacitor, consuming more power. The maximum signal is determined by the linear range of the transconductor. This work showed that both the noise and the maximum signal can be amplified by including a capacitive attenuator in the feedback path of filter.

In order to increase the dynamic range, I explored the non-linear operation of the filters, including jump resonance. Unlike harmonic distortion and gain compression which slowly increase with the input amplitude, jump resonance is not present in a linear system, but develops in the presence of strong nonlinearity. It is characterized by a discontinuous jump in the frequency response near the resonant peak. I have analyzed the behavior using both describing function and state-space techniques. Then, I developed a novel graphical analysis technique. Finally, I design, built, and tested a circuit for avoiding jump resonance for audio filters.

Finally, I took advantage of nonlinearities in a filtering system to build a micropower companding speech processor. This system implements the companding speech processing algorithm to improve speech comprehension in moderate noise environments. The sixteen channel system increases the spectral contrast of speech signals by performing an adjustable two-tone suppression function, replacing the function of a normally function cochlea for hearing aid or cochlear implant users. The system runs on less than 60uW of power, a consumption so low it could run for 6 months on a standard hearing aid battery.

Thesis Supervisor: Rahul Sarpeshkar
Title: Associate Professor of Electrical Engineering

Table of Contents

Table of Figures

8

9

11

Acknowledgements

Biographical Note

Christopher Salthouse received his bachelor's and master's degrees in Electrical Engineering from the Massachusetts Institute of Technology in 2000. His master's thesis, *Improvements to the Non-Intrusive Load Monitor*, was supervised by Prof. Steven Leeb in the Laboratory of Electromagnetic and Electronic Systems at M.I.T. He has worked for the last six years for Prof. Rahul Sarpeshkar in the Research Laboratory of Electronics at M.I.T. He has also been on the teaching staff for 6.002 Introduction to Electronic Circuits and 6.973 Analog VLSI and Biological Systems. He has presented his work at the International Symposium on Circuits and Systems and in the Journal of Solid State Circuits and the IEEE Transactions on Circuits and Systems I.

Chapter 1 : Introduction

Filters are one of the basic building blocks of analog circuits. For decades, researchers have studied electronic filters [1]. So much is understood about filter design that one popular introductory textbook tells readers, "Filter design is one of the very few areas of engineering for which a complete design theory exists, starting from specification and ending with a circuit realization [2]." While this quote may be true for linear filter design, it remains untrue for the non-linear regime. For small signals, filters can be approximated by linear models, but as signals grow this approximation weakens. The range of signals for which the approximation holds is the dynamic range. It is the range, measured in dB, between the smallest signal the system can process and the largest. Every filter has some signal that is so small that it cannot be distinguished from noise and some signal so large that the filter does not perform properly. This thesis is about the second half of that sentence, defining the high end of analog filters for audio applications.

This thesis is divided into five chapters. The first chapter provides a background of technologies used in analog filtering in the audio frequency range. The second chapter presents my work on the capacitive-attenuation technique for increasing the linear range of gm-c filters. The third chapter explores jump resonance, a non-linear behavior of filters, including my novel graphical analysis method and adaptive circuit solution. The fourth chapter describes the application of micropower analog filters to my companding speech processor to improve speech recognition for the hearing impaired in moderate noise environments.

Section 1. Audio Filter Applications

There are a variety of applications for audio filtering including: automatic speech recognition, speech enhancement, speech coding, and processing for cochlear implants. In each of these applications, the filters are not required to be linear, but must demonstrate some of the properties that are guaranteed by linearity. So, linearity is often used as a proxy for the properties that are needed. So, I will first analyze the requirements for linearity and then look at the places where these requirements can be relaxed for the applications of interest.

Automatic speech recognition has been the topic of much research in the last half century. A general architecture has emerged. Phonemes are extracted and then the sequence of phonemes is interpreted. Because phonemes are defined by spectral features such as formants, the phoneme detection stage involves filtering as either an explicit or implied process. In some designs, the speech is separated into frequency bands by a bank of bandpass filters. Formants are then extracted by measuring the power in each band. Alternatively, the two processes can be combined in the digital domain using a Fast Fourier Transform (FFT) or Linear Predictive Coding (LPC). LPC finds a function that predicts the next sample value based on a series of previous samples. Because this is equivalent to defining a discrete time filter for the channel, the LPC is often accomplished by first transforming the samples to the frequency domain, although some algorithms function solely in the time domain [3]. After the phonemes are extracted, they are matched to possible words. Then the possible words are weighted based on sentence level probabilities. In these systems, the phoneme identification may be amenable to analog signal processing while the later classification steps are probably more appropriate for digital computation.

One of the most visible areas of research in audio processing has been perceptual coding. The factor of 10 compression achieved in MP3 encoding is in large part due to these encoding techniques that perform selective compression based on the ability of humans to hear the artifacts. The encoding process can be divided into two stages: perceptual modeling and variable compression. In the perceptual modeling stage, a frequency analysis is performed on a window of the sound and a masking profile is created. This masking profile is then used to set the maximum errors when each frequency band is compressed [4-5]. This masking profile can be developed using special analog filters, while the variable compression is performed with digital computation.

Speech enhancement is a related area that has shown some success and promises even more. Speech enhancement is the general term for a variety of techniques used to improve the intelligibility of speech for either human or computer listeners. The basis of these methods is amplifying those parts of the signal, as defined in time and frequency, which have a high signal to noise ratio while attenuating those portions where the noise dominates. One of the simpler algorithms performs non-linear expansion based on signal amplitude in each frequency band

[6]. More advanced techniques use adaptive filtering such as Kalman filters to maximize the signal to noise ratio based on noise estimates measured during pauses in speech signals [7]. Spectral Subtraction lies somewhere between these systems in complexity, with frequency domain subtraction of the noise signal from the composite signal [8]. The performance of these systems is often characterized by resulting SNR, but work has also been done comparing the automatic speech recognition scores on the output of these systems [8]. These systems have been implemented in both DSP and analog systems.

While the above applications are aimed at a broad audience, the hearing impaired can benefit even more from audio frequency filters. It is estimated that one million Americans are sensorineurally deaf. They can only hear with the aid of a cochlear implant, a biomedical device that directly stimulates the nerves in the cochlea in response to sound [9]. A microphone captures the sound. Then the dynamic range is decreased using an automatic gain control. The sound is divided in to frequency bands. The power in those bands is calculated. A second stage of compression is applied. Finally, the neurons are stimulated [10-11]. Active research on cochlear implants is aimed at optimizing the compression strategy and minimizing the power consumption to extend battery life.

Section 2. Analog Filter Topologies

In any of these applications, micropower analog filters are generally built in one of two technologies: Gm-C and log domain. Both of these technologies include a variety of topologies. The Gm-C domain is defined as filters built using linear voltage to current converters, the Gm, and capacitors, the C. These filters are also frequently referred to as OTA-C filters. Log domain filters differ in that they store state on capacitors which contain a log representation of the signal and use exponential voltage to current converters. They go by many names including: Exponential State Space (ESS), Dynamic Translinear (DTL), and Bernoulli cells.

A: Gm-C

Sanchez-Sinencio and Silva-Martinez provided an excellent overview of Gm-C filters in greater depth than presented here [12]. Research in Gm-C filters addresses both the transconductor

design and the filter topology. A survey of the options available in each and the tradeoffs associated with them is presented here.

Single Ended

Practical transconductors are either single transistor input transconductors or differential pair transconductors. Single transistor input transconductors can be built using saturation region MOSFETs, triode region MOSFETs, or forward active region bipolar transistors. The saturation MOSFETs and active bipolar transistors have the advantage of yielding the simplest possible circuits. Each transconductor is one transistor. The disadvantage of this design is the limited linear range. In fact, it is not linear at all as shown in Fig. 1.1. In subthreshold operation, the transistor curve is exponential. Above threshold, it is a square law. For small signal swings, it can be approximated by a linear function [13].
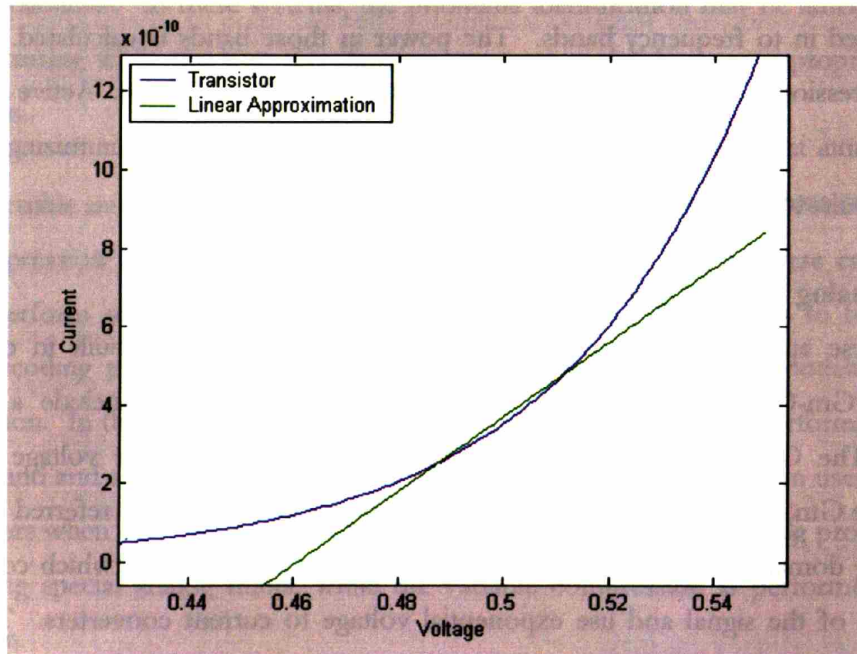
Figure 1.1: Linear Approximation of a Single Saturated Transistor

Triode transconductors offer improved linearity. One commonly used approximation for the current in a MOSFET when the gate-source potential is strong enough to create strong inversion, but the drain-source voltage is small enough that the device is not is saturation is given by Eq. 1.1 [14]. $I_{DS} = \dfrac{W}{L} \mu C_{ox} \left[(V_{GS} - V_{TS}) V_{DS} - \dfrac{\alpha}{2} V_{DS}^2 \right]$ (1.1)

20

If α is small, the drain current is proportional to the drain-source voltage with the proportionality defined by the gate voltage. In fact, this technique has been used extensively [15-17]. Sample curves from simulation are shown in Fig. 1.2. The linear range at the bottom is limited by the threshold voltage and at the top by the input signal. Because this is based on above-threshold transistors, the current levels are difficult to achieve below the order of microAmps. Unlike the saturation mode transconductor, the triode mode transconductor requires additional circuitry to maintain the operating condition as shown in Fig. 1.3.
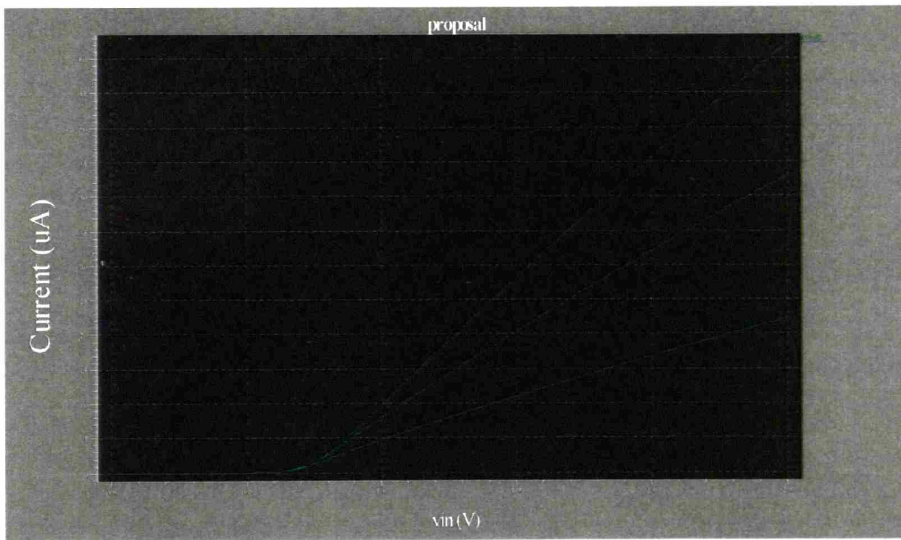
Figure 1.2: Triode Region Transconductor Simulations
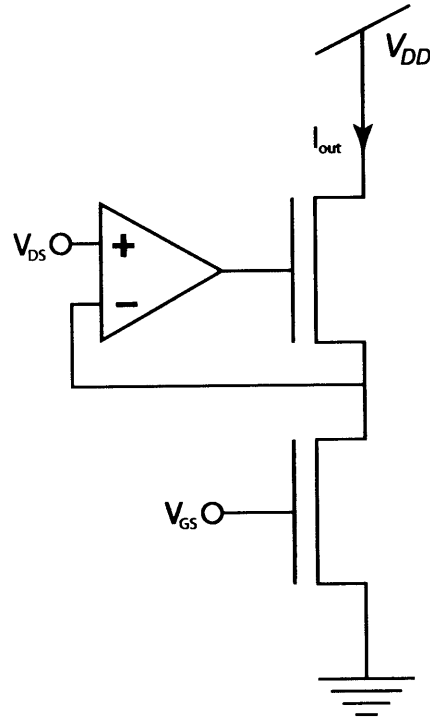
Figure 1.3: Triode Transconductor Circuit

Differential

Differential transconductors are the other broad class of transconductor circuits. The simplest implementation is shown in Fig. 1.4. Transistors M_1 and M_2 split the current, I_{bias}, between the two legs. The current mirror formed by M_3 and M_4 performs the current subtraction to create I_{out}. The derivation of I_{out} for subthreshold transistors is straightforward.
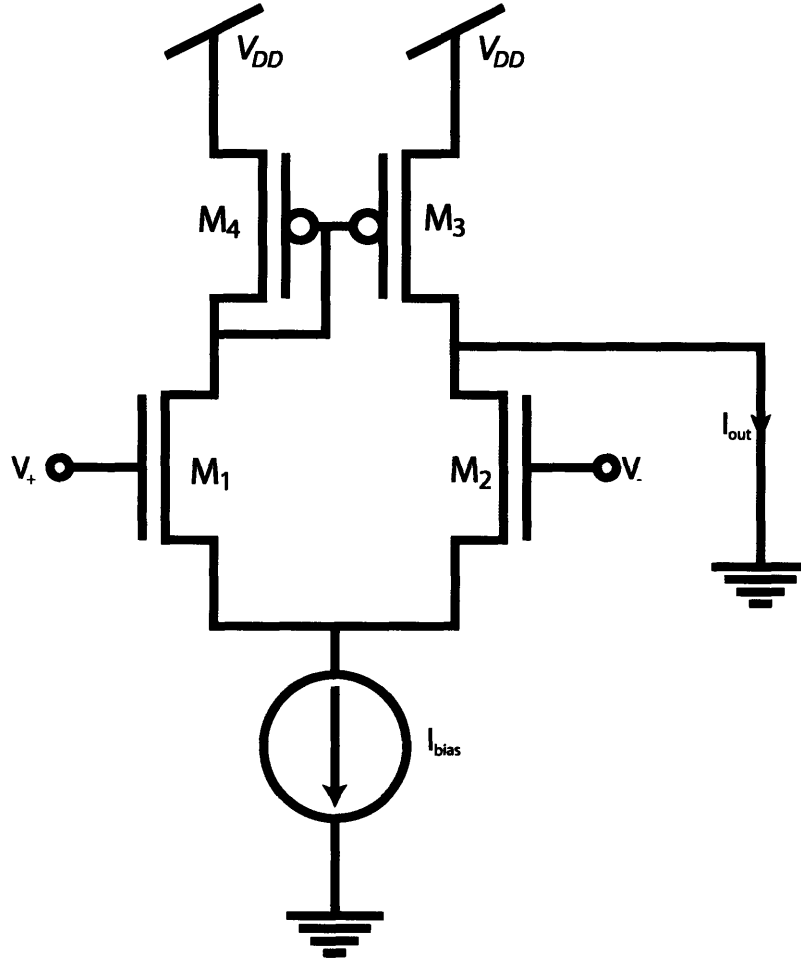
Figure 1.4: Differential Transconductor Circuit

$$I_1 = I_s e^{\frac{\kappa V_+ - V_s}{\phi_t}} \; ; I_2 = I_s e^{\frac{\kappa V_- - V_s}{\phi_t}} \tag{1.2}$$

$$I_{out} = I_1 - I_2 = I_s \left(e^{\frac{\kappa V_+ - V_s}{\phi_t}} - e^{\frac{\kappa V_- - V_s}{\phi_t}} \right) \tag{1.3}$$

$$I_{bias} = I_1 + I_2 = I_s \left(e^{\frac{\kappa V_+ - V_s}{\phi_t}} + e^{\frac{\kappa V_- - V_s}{\phi_t}} \right) \tag{1.4}$$

$$I_{out} = I_{bias} \left(\frac{e^{\frac{\kappa V_+}{\phi_t}} - e^{\frac{\kappa V_-}{\phi_t}}}{e^{\frac{\kappa V_+}{\phi_t}} + e^{\frac{\kappa V_-}{\phi_t}}} \right) \tag{1.5}$$

$$I_{out} = I_{bias} \tanh \left(\frac{\kappa}{\phi_t} \left(\frac{V_+ - V_-}{2} \right) \right) \tag{1.6}$$

23

The linear range is $2\phi_t/\kappa$, approximately 75mV.

The advantages of the differential pair are inherent differential operation and programming to very low transconductance. The differential operation ensures common mode rejection and cancellation of noise from bias currents. The low transconductance is achieved by programming the bias current as opposed to a DC gate bias for the single ended designs. Both of these properties can be compared to a pseudo-differential transconductor built by combining two singled ended transconductors and subtracting the output current of the inverting input [18]. This transconductor requires matching of two transconductances and the transconductors will continue to respond to the common mode requiring increased current.

Linear range is one of the major limitations with any of these saturation mode topologies, either single ended or differential. A wide variety of techniques have been used to improve this linear range, but most work has focused on attenuation and degeneration [12, 19].
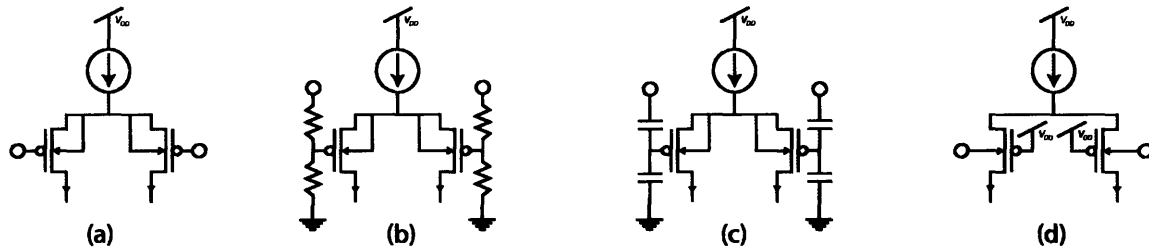
Figure 1.5: Attenuation Techniques For Differential Transconductors

Attenuation is the simplest of the techniques; the signal is simply scaled by a factor less than 1 prior to controlling the differential pair. Figure 1.5 shows a simple differential pair (a) and three attenuation methods (b-d). Resistive attenuation (1.5b) has been used in discrete OTAs, but is impractical for integrated circuits because of power and space limitations. A capacitive attenuator (1.5c) can be used for integrated circuits, but a DC path must be provided as discussed in Chapter 2. The final circuit (1.5d) is also an attenuation scheme, but with intrinsic attenuation. By using the well as the input to the circuit rather than the gate, the transconductance of the transistor is decreased. This is classified as an attenuation approach

24

because it is a feed-forward approach that is exactly equivalent to supplying a smaller signal. Physically, the classification of this approach as an attenuation scheme is justified by recognizing the surface potential of the transistor as the control terminal with capacitances from gate and well controlling this terminal [19].
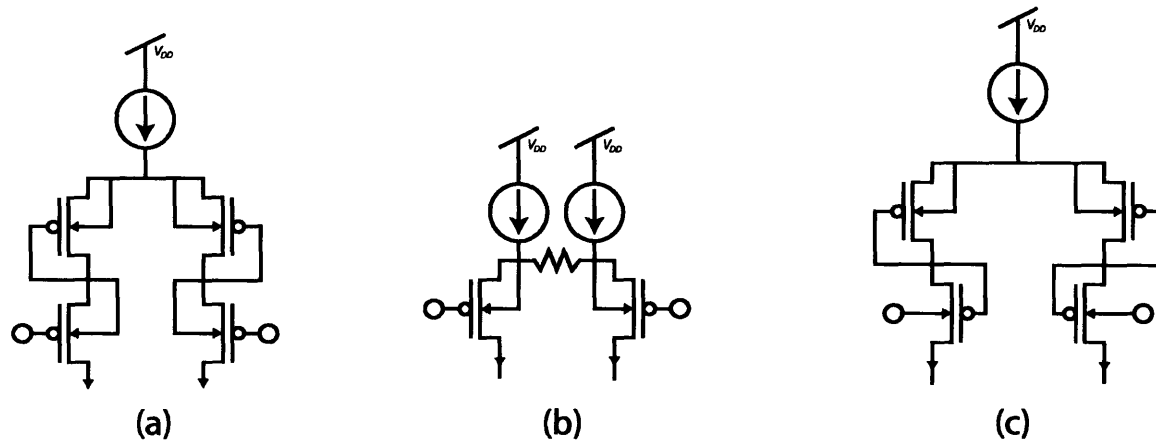
Figure 1.6: Degeneration Techniques for Differential Transconductors

Degeneration schemes also lower the voltage across the control terminal, but they do it through feedback. Figure 1.6 shows three techniques for degeneration. In all three, the current at the output passes through the degenerating device creating a voltage that subtracts from the input voltage applied to the transconducting transistors. The circuit in part (a) has diode degeneration. The voltage across the diode connected transistor lowers the voltage on the source of the input PMOS, decreasing the current through the device. If the diode and transconducting devices are the same dimensions, they will split the input voltage evenly, doubling the linear range. This technique limits the common mode range of the circuit because the diodes are continually dropping voltage across them.

Because it is only the differential mode where degeneration is required, circuits such as the one shown in part (b) are often used. No common mode current flows through the degenerating device, but any imbalance between the two legs is fed through the degenerating device creating a voltage imbalance that functions as in part (a).

25

The final example of degeneration presented here (c) is a combination of source and gate degeneration. The voltage induced across the diode connected transistor lowers the voltage on both the source and the gate. Both terminals act to lower the current [19].

B: Log-Domain Filters

Unlike transconductance filters, log domain filters do not approximate the transistor as a linear element. Instead, the transistor is used as an exponential device. This is demonstrated for a bipolar transistor and a subthreshold MOSFET in Fig. 1.7 and Eq. 1.7 and 1.8.
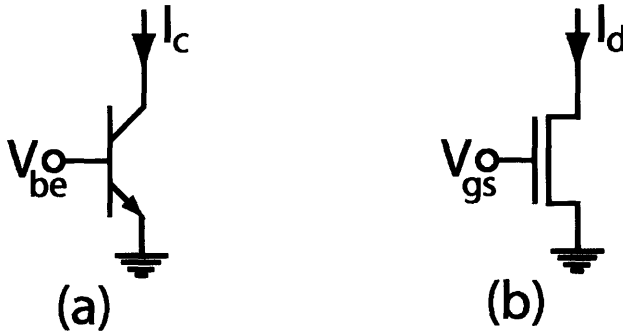
Figure 1.7: Exponential Devices: (a) is described by Eq. 7. (b) is described by Eq. 8.

$$I_c = I_s e^{\frac{V_{be}}{\phi_t}} \tag{1.7}$$

$$I_d = I_{os} e^{\frac{\kappa V_{gs}}{\phi_t}} \tag{1.8}$$

As with transconductor filters, there are a variety of design methods that are used to generate log domain filters. Three of the most published techniques are associated with the keywords Exponential State Space (ESS), Dynamic Trans Linear (DTL), and Bernoulli Cell.

Exponential State Space

Much of the excitement about log domain filtering began with a series of papers by Frey on what he termed Exponential State Space filters [20]. The ESS method begins with a state space description of the desired filter such as the second order system in Eq. 1.9.

$$\begin{vmatrix} \dot{x}_1 \\ \dot{x}_2 \end{vmatrix} = \begin{vmatrix} -\omega_0/Q & -\omega_0 \\ \omega_0 & 0 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} + \begin{vmatrix} \omega_0 \\ 0 \end{vmatrix} U$$

$$Y = \begin{vmatrix} 1/Q & 0 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix}$$

(1.9)

Then the exponential mapping, Eq. 1.10, is applied.

$$f_i(V_i) = x_i$$

(1.10)

This transforms the generic state space description to the form in Eq. 1.11.

$$F'(V)\dot{V} = AF(V) + BU$$
$$Y = CF(V) + DU$$

(1.11)

Using the bipolar equation (Eq. 1.7) to transform the filter equations (Eq. 1.9) gives Eq. 1.12.

$$\dot{v}_1 \frac{I_1}{\phi_t} = -\frac{\omega_0}{Q} I_1 - \omega_0 I_2 + \omega_0 U$$

$$\dot{v}_2 \frac{I_2}{\phi_t} = \omega_0 I_1$$

(1.12)

$$C_1 \dot{v}_1 = -\frac{C_1 \omega_0 \phi_t}{Q} - C_1 \omega_0 \phi_t \frac{I_2}{I_1} + \frac{C_1 \omega_0 \phi_t}{I_1} U$$

$$C_2 \dot{v}_2 = C_2 \omega_0 \phi_t \frac{I_1}{I_2}$$

(1.13)

In this form, it can be seen that the two equations define currents onto capacitors C_1 and C_2.

The first term is a constant current source; in general the terms involve division of currents. Because the currents are derived from an exponential function of voltage, this division can be accomplished by subtraction of voltage. Some of the building block circuits used by Frey are shown in Fig. 1.8. The first two use a combination of n-type and p-type transistors while the third has been proposed for higher frequency applications since it only requires n-type transistors [19-20].

27
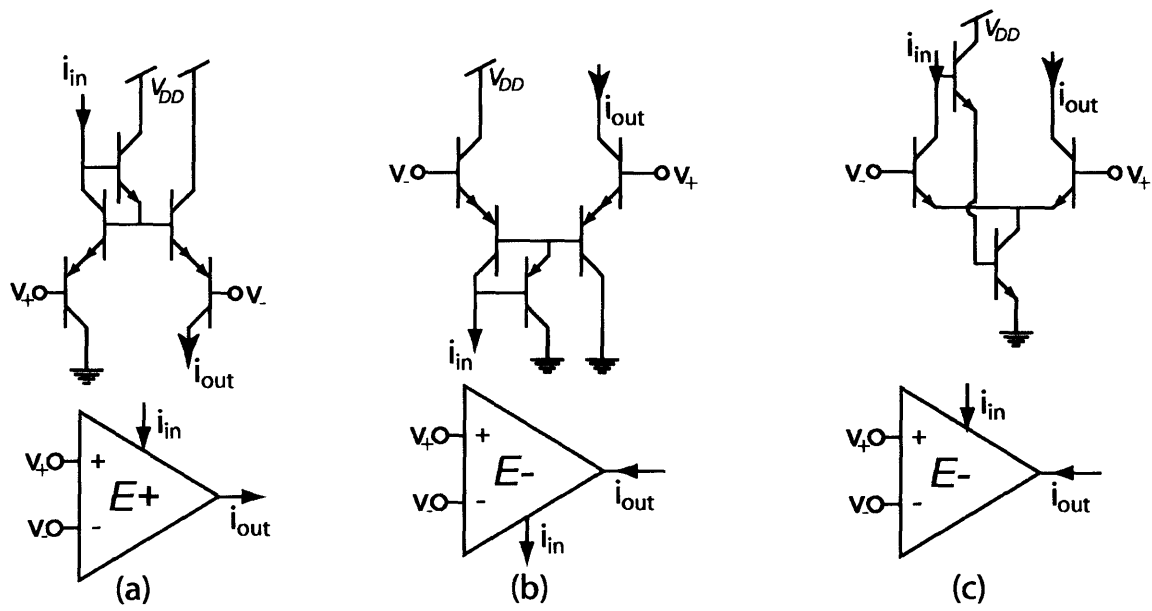
Figure 1.8: Frey's ESS Units
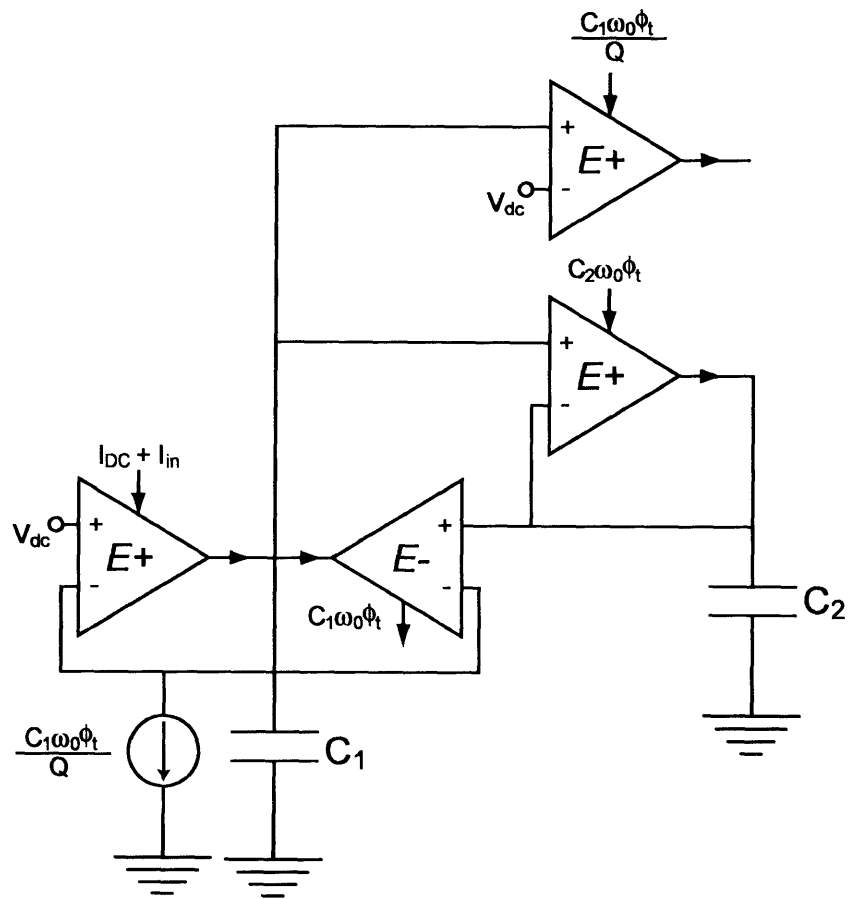
Figure 1.9: ESS Filter

These blocks are then used to build up the circuit for the filter as shown in Fig. 1.9. In practice, this circuit will not work because of the DC operating point. C_2 is only connected to a current source with no current sinks. The operating point can be adjusted by adding an additional input to the state space representation. Because the filter is linear, a DC value applied to this input will simply shift the output. Equation 1.9 becomes

$$\begin{vmatrix} \dot{x}_1 \\ \dot{x}_2 \end{vmatrix} = \begin{vmatrix} -\omega_0/Q & -\omega_0 \\ \omega_0 & 0 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix} + \begin{vmatrix} \omega_0 \\ 0 \end{vmatrix} U + \begin{vmatrix} 0 \\ \omega_0 \end{vmatrix} U_{DC}$$

$$Y = \begin{vmatrix} 1/Q & 0 \end{vmatrix} \begin{vmatrix} x_1 \\ x_2 \end{vmatrix}$$

(1.14)

Any negative value of U_{DC} will give a viable operating point, but it is straightforward to solve for the required input current to set x_1 and x_2 to the same value. If the primary input offset is I_{DC}, then the DC input should be $\dfrac{I_{DC}Q}{Q+1}$. Adding this to the circuit creates the circuit in Fig. 1.10.

Figure 1.10: Improved ESS Filter
Dynamic Trans Linear

The Dynamic Trans Linear (DTL) technique views log domain circuits as a generalization of translinear circuits. The basic building block is shown in Fig. 1.11 [22]. This block can be simply analyzed.

$$I_q = I_s e^{\frac{V}{\phi_t}} \tag{1.15}$$

$$I_{cap} = C\dot{V} \tag{1.16}$$

$$\dot{I}_q = \frac{\dot{V}}{\phi_t} I_s e^{\frac{V}{\phi_t}} = \frac{\dot{V}}{\phi_t} I_q \tag{1.17}$$

30

$$\frac{\dot{I}_q}{I_q} = \frac{I_{cap}}{C\phi_t}$$

(1.18)

Figure 1.11: DTL Building Block

Equation 1.18 is the standard equation used in this method. The state space description in Eq. 1.14 can be represented in the current domain as Eq. 19.

$$\left|\begin{matrix} \dot{I}_1 \\ \dot{I}_2 \end{matrix}\right| = \left|\begin{matrix} -\omega_0/Q & -\omega_0 \\ \omega_0 & 0 \end{matrix}\right| \left|\begin{matrix} I_1 \\ I_2 \end{matrix}\right| + \left|\begin{matrix} \omega_0 \\ 0 \end{matrix}\right| I_U + \left|\begin{matrix} 0 \\ \omega_0 \end{matrix}\right| I_{DC}$$

$$Y = \left|\begin{matrix} 1/Q & 0 \end{matrix}\right| \left|\begin{matrix} I_1 \\ I_2 \end{matrix}\right|$$

(1.19)

The current on to capacitors C_1 and C_2 can be solved by manipulating Eq. 1.19.

$$I_{C_1} = C_1 \phi_t \frac{\dot{I}_1}{I_1} = C_1 \phi_t \left(-\left(\frac{\omega_0}{Q}\right) - \frac{I_2 \omega_0}{I_1} + \frac{I_U}{I_1} \right)$$

$$I_{C_2} = C_2 \phi_t \frac{\dot{I}_2}{I_2} = C_2 \phi_t \left(\frac{I_1 \omega_0}{I_2} - \frac{I_{DC}}{I_2} \right)$$

(1.20)

These equations can be implemented using Gilbert multipliers as shown in Fig. 1.12.
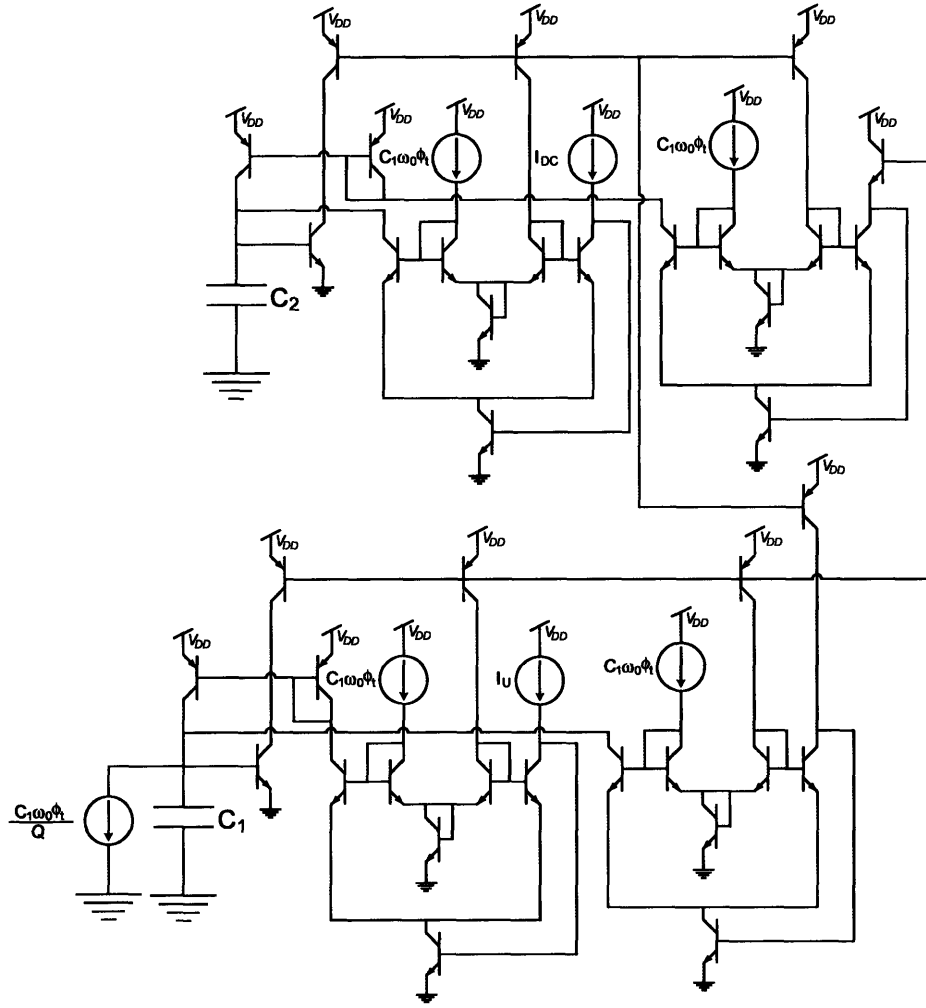
Figure 1.12: DTL Filter

The transistor count can be reduced by factoring the multiplications in Eq. 1.20 to write them as Eq. 1.21

$$I_{C_1} = -\left(\frac{C_1\phi_t\omega_0}{Q}\right) + \frac{C_1\phi_t\omega_0}{I_1}\left(-I_2 + \frac{I_U}{\omega_0}\right)$$

$$I_{C_2} = C_2\phi_t\frac{I_2}{I_2} = C_2\phi_t\omega_0\left(I_1 - \frac{I_{DC}}{\omega_0}\right)$$

(1.21)

The two multipliers that feed C_1 are partially combined, as are the two multipliers that feed C_2. This creates the circuit in Fig. 1.13.
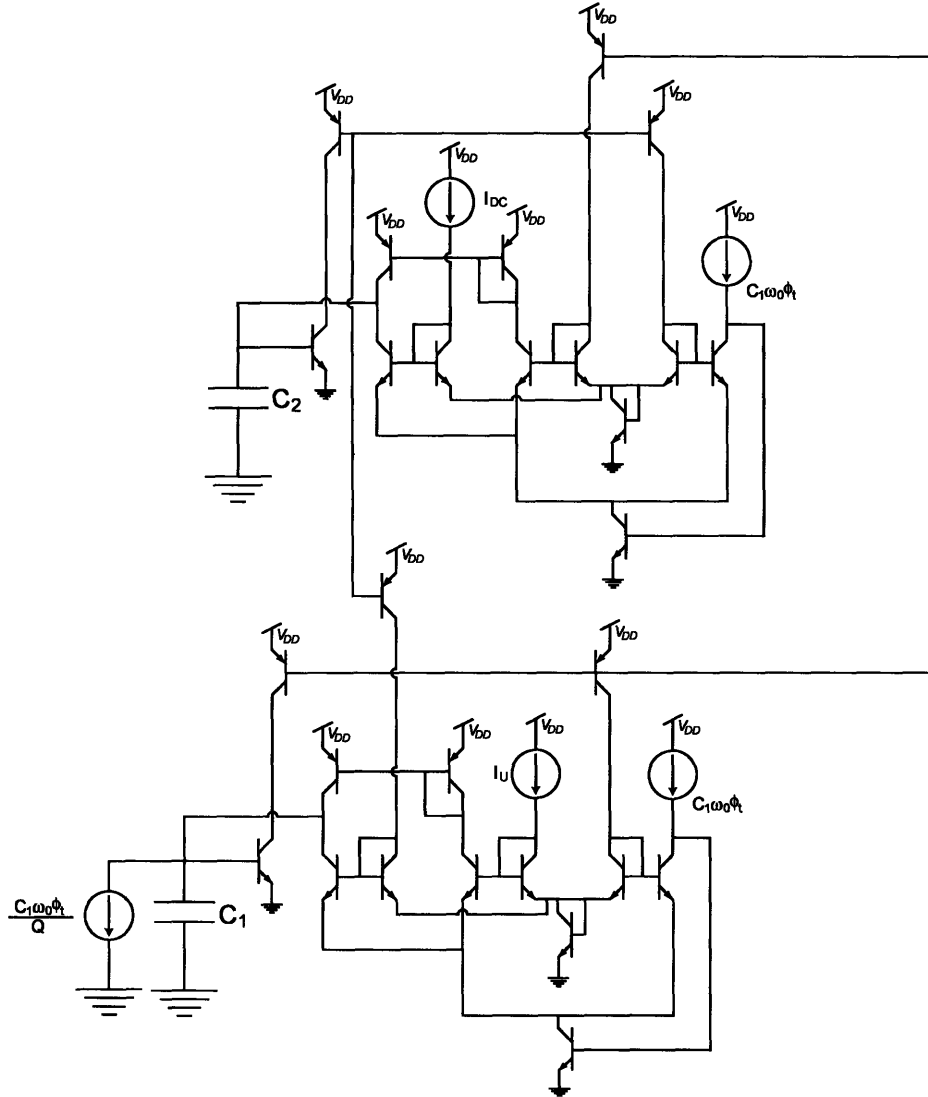
Figure 1.13: Improved DTL Filter

The DTL technique is primarily an analysis technique and applies to any circuit where the linear state variable is a current stored as a base-emitter voltage, V_{be}, on a capacitor. So, other DTL circuits could also be built to realize this system response. But, this design is a useful example for comparing the different techniques, as is done following the Bernoulli Cell technique.

Bernoulli Cell

The Bernoulli Cell based transconductors are based on the building block shown in Fig. 1.14. Quite simply, a capacitor is placed at the emitter of a transistor. This is remarkably similar to the building block used for DTL circuits, where the capacitor is placed on the base of the transistor. In fact, practical Bernoulli Cells circuits can be analyzed using the DTL principle.

The DTL analysis cannot be applied to this building block cell, however, because the voltage difference between the base and emitter of the transistor are varying. In comparing the building block to the DTL building block, the two primary differences are the negative sign in the exponent because of control of the transistor through the emitter, and the inclusion of transistor current in the capacitor current, I_{cap}.
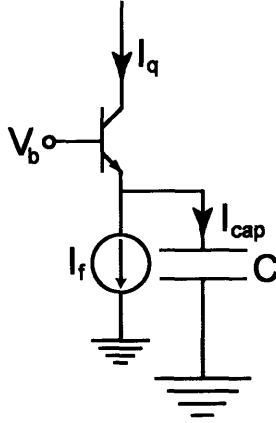
Figure 1.14: Bernoulli Cell Building Block

$$I_q = I_s e^{\frac{V_b - V_c}{\phi_t}}$$
(1.22)

$$\frac{\dot{I}_q}{I_q} = -\frac{\dot{V}_c}{\phi_t}$$
(1.23)

$$I_{cap} = I_q - I_f$$
(1.24)

$$C\dot{V}_c = I_{cap}$$
(1.25)

$$\frac{\dot{I}_q}{I_q} = -\frac{I_{cap}}{C\phi_t}$$
(1.26)

$$-\frac{\dot{I}_q C\phi_t}{I_q} = I_q - I_f$$
(1.27)

$$\dot{I}_q C\phi_t = I_f I_q - I_q^2$$
(1.28)

Following the same flow of analysis that was used for the DTL system in Eq. 1.15 through 1.20, it is clear that this system is different. This circuit is called a Bernoulli Cell because this equation is of the Bernoulli form [23]. Mathematicians suggest that this circuit can be linearized by making the substitution in Eq. 1.29.

$$T = \frac{1}{I_q} \qquad (1.29)$$

Circuit designers should recognize this substitution as the conversion from emitter voltage to base voltage. The negative sign in the exponent represents this reciprocal. The substitution continues with Eq. 1.30-1.32.

$$\dot{T} = -\frac{\dot{I}_q}{I_q^2} \qquad (1.30)$$

$$\frac{\dot{T}}{T^2} C\phi_t = \frac{1}{T^2} - \frac{I_f}{T} \qquad (1.31)$$

$$\dot{T} C\phi_t = 1 - T I_f \qquad (1.32)$$

The form of Eq. 1.32 shows that this structure has less flexibility than the techniques presented earlier. It is easier to see this limitation if a filter is built using this building block. The simplest filter uses just one of these building blocks. Even with this constraint, multiple filters may be built two are presented in Fig. 1.15. Circuit A is from [23] and Circuit B is from [24].
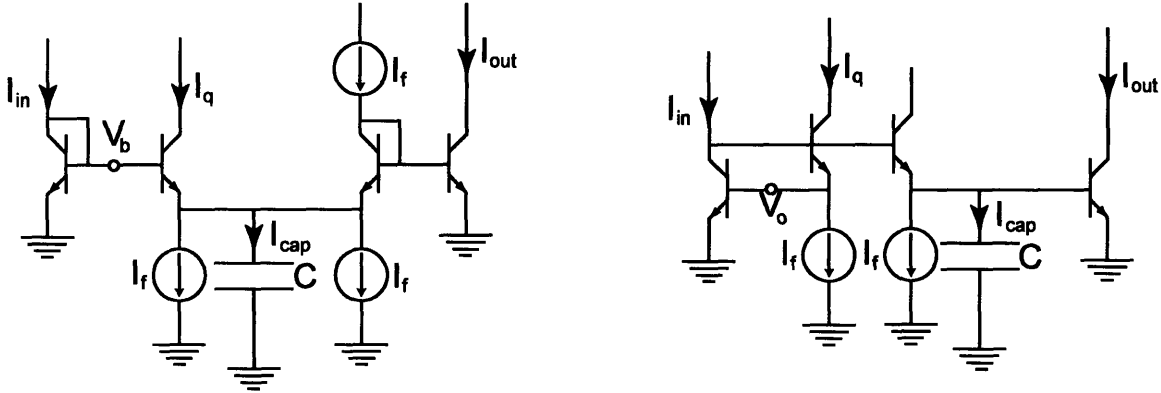
Figure 1.15: Bernoulli Cell Filters

I_{out} as a function of I_{in} is derived in Eq. 33-44. The translinear principle, called the static translinear principle by the DTL practitioners, is used to generate Eq. 1.33.

$$I_{out} = \frac{I_{in} I_f}{I_q} \qquad (1.33)$$

35

The transistor relation is then used to generate

$$I_q = I_s e^{\frac{V_b - V_c}{\phi_t}}$$ (1.34)

And the capacitor relation is used.

$$C\dot{V}_c = I_q - I_f$$ (1.35)

$$\phi_t \frac{\dot{I}_q}{I_q} = \dot{V}_b - \dot{V}_c$$ (1.36)

$$\dot{V}_b = \phi_t \frac{\dot{I}_{in}}{I_{in}}$$ (1.37)

$$\phi_t \frac{\dot{I}_q}{I_q} = \phi_t \frac{\dot{I}_{in}}{I_{in}} - \dot{V}_c$$ (1.38)

$$C\phi_t \frac{\dot{I}_q}{I_q} = C\phi_t \frac{\dot{I}_{in}}{I_{in}} + I_f - I_q$$ (1.39)

Then Eq. 1.33 is substitute into Eq. 1.39.

$$\dot{I}_q = \frac{\dot{I}_{in} I_f}{I_{out}} - \frac{I_{in} I_f \dot{I}_{out}}{I_{out}^2}$$ (1.40)

$$C\phi_t \left(\frac{\frac{\dot{I}_{in} I_f}{I_{out}} - \frac{I_{in} I_f \dot{I}_{out}}{I_{out}^2}}{\frac{I_{in} I_f}{I_{out}}} \right) = C\phi_t \frac{\dot{I}_{in}}{I_{in}} + I_f - \frac{I_{in} I_f}{I_{out}}$$ (1.41)

$$C\phi_t \left(\frac{\dot{I}_{in} - \frac{I_{in} \dot{I}_{out}}{I_{out}}}{I_{in}} \right) = C\phi_t \frac{\dot{I}_{in}}{I_{in}} + I_f - \frac{I_{in} I_f}{I_{out}}$$ (1.42)

$$C\phi_t \dot{I}_{out} = I_{out} I_f - I_{in} I_f$$ (1.43)

$$\frac{I_{out}}{I_{in}} = \frac{I_f}{I_f + sC\phi_t}$$ (1.44)

With a complete circuit as in Fig. 1.15, it is possible to use the DTL principle for analysis. The network that linearizes the Bernoulli Cell may also be viewed as creating a DTL circuit. The

36

base of the output transistor is driven by a level shifted version of the capacitor current. From Eq. 1.18 in the DTL section,

$$\frac{\dot{I}_{out}}{I_{out}} = \frac{I_{cap}}{c\phi_t} \qquad (1.45)$$

Substituting in for I_{cap} yields

$$\frac{\dot{I}_{out}}{I_{out}} = \frac{I_q - I_f}{c\phi_t} \qquad (1.46)$$

Equation 1.33 can be used to remove I_q from the equation.

$$\frac{\dot{I}_{out}}{I_{out}} = \frac{I_{in}I_f}{c\phi_t I_{out}} - \frac{I_f}{c\phi_t} \qquad (1.47)$$

$$\dot{I}_{out}c\phi_t = I_{in}I_f - I_f I_{out} \qquad (1.48)$$

$$\frac{I_{out}}{I_{in}} = \frac{I_f}{I_f + sC\phi_t} \qquad (1.49)$$

Log Domain Filter Comparison

From Eq. 1.49 the loss of generality in the Bernoulli Cell method as compared to ESS and DTL is clear. This building block is simply a first order low pass filter. These blocks can be combined to make general purpose filters, but the added overhead can be significant. Figures 1.10 and 1.13 show that the ESS and DTL methods create similar circuits. Since the Bernoulli cell is a subset of DTL, it is simply an option when the circuit being built fits the requisite form.

Class-AB Operation of Log Domain Filters

In the log domain circuits presented so far, the signal can not exceed the bias current, but log domain circuits can be operated in a class-AB mode to increase the maximum signal [25]. This maximum amplitude is similar to the transconductance requirement that transconductors be biased with a current equal to their maximum current. But, the limitations for the linear range are very different between Gm-C and log domain circuits. For Gm-C filters, as the signal grows it begins to clip on both the top and the bottom as shown in part (c) of Fig. 1.16, but with log domain circuits the signals only clip on the bottom as in part (b). Signals can be processed without clipping even if they are much larger than the bias current, as long as they never become negative.

37
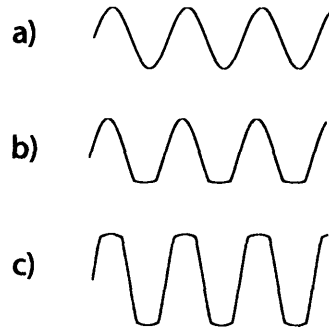
Figure 1.16: Clipping

This property of log-domain filters allows the creation of a special type of differential filters. As with other differential systems, the composite variable is the difference of the signals in two paths. But, rather than keeping the sum of the two constant, a rule is created such that both currents are always positive. A common rule is that the product of the two variables is constant. Figure 1.17 demonstrates the difference. Part (a) shows the differential signals in a constant common mode differential scheme. Part (b) shows the differential signals used in a class-AB scheme.
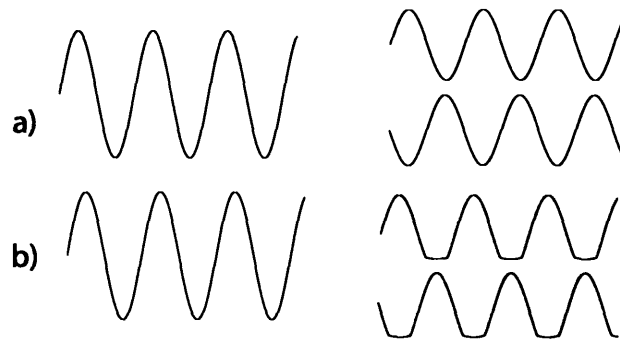
Figure 1.17: Differential Signals: a) Constant Common Mode. b) Class AB

This technique works for the same reason as other differential techniques. Any common mode effects are canceled out when passed through identical linear systems. So, a particularly useful common mode is chosen for this application. Whereas a constant common mode approach doubles the linear range, since the maximum swing of two channels are now added to each other, the class-AB approach can increase the linear range by much larger amounts.

Ultimately it is limited by the assumptions of identical linear systems and the ability of the input circuit to create the signals. This distortion is well modeled by the distortion in a simple current

mirror. As the current increases, some of the current is used to charge up the gate capacitance. As the current decreases, this capacitance is discharged through the transconductance of the input device. If this current mirror is driven by a square wave, the output will not follow the falling edge. The distortion introduced by this pole is shown in Fig. 1.18 which limits the modulation index to approximately a factor of 5.
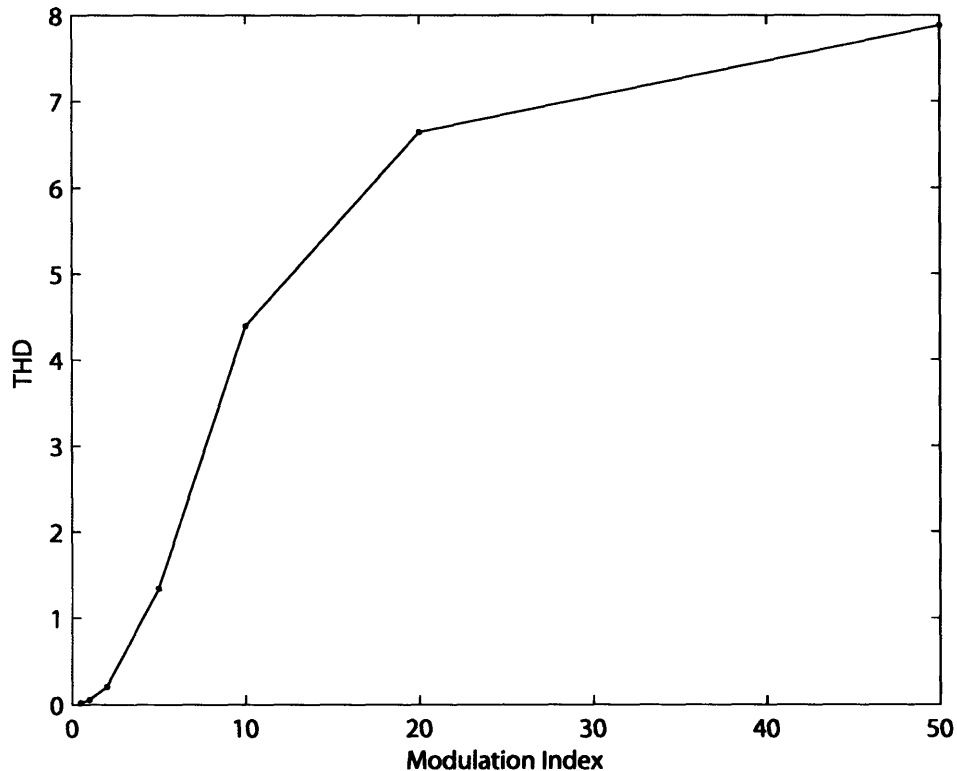
Figure 1.18: The signal distortion introduced by the input circuit increases with modulation index, the ratio of the signal maximum to the geometric mean of the channels.

Section 3. Standard Dynamic Range Measurement Techniques

The high end of dynamic range of filters is defined differently for different applications. For generic filter design, total harmonic distortion (THD) is the most often used measure [26-28]. But, the THD level that defines the top of the dynamic range also differs with little explanation as to the chosen percentage. In communications systems, other standards are often used [29-

31]. Intermodulation distortion (IMD) is designed as a test of how the system performs with real world signals. More abstract measures include compression and intercept points.

A: Total Harmonic Distortion

Harmonic distortion is the generation of signals at harmonics of the input. Non-linear functions generate these harmonics as shown in Eq. 45- 48. Given the input signal

$$x(t) = A \cos \omega t \tag{1.50}$$

and a system described by the function

$$y(t) = \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) \tag{1.51}$$

the output is

$$y(t) = \alpha_1 A \cos \omega t + \alpha_2 A^2 \cos^2 \omega t + \alpha_3 A^3 \cos^3 \omega t \tag{1.52}$$

This can be simplified to

$$y(t) = \frac{\alpha_2 A^2}{2} + \left(\alpha_1 A + \frac{3\alpha_3 A^3}{4} \right) \cos \omega t$$
$$+ \frac{\alpha_2 A^2}{2} \cos 2\omega t + \frac{\alpha_3 A^3}{4} \cos 3\omega t \tag{1.53}$$

Total Harmonic Distortion (THD) is just the ratio of the power in all of the harmonics versus the fundamental [31]. There are two problems with this approach. The first is that the harmonics are filtered in later stages. So the system may display other non-linear qualities, such as gain compression discussed below, while not having significant harmonics. The other problem is that harmonic distortion becomes imprecise for systems that are moderately non-linear. A square wave has only 11% third harmonic distortion as measured in power.

B: Compression Point

An alternative to harmonics is to look at the gain of the system. Most systems begin to saturate as they become non-linear as shown in Fig. 1.16. In differential designs that rely upon common mode features canceling out, such as the class-AB topology, compression can introduce differential artifacts.

40

Section 4. The Auditory System

In building audio processing systems, it is natural to look at how the human auditory system functions as a reference design. The design is not being copied as a whole, so key features may be lost. For instance, non-linearity in the auditory system may be acceptable because of the massive parallel nature of the auditory nerve. But, it also may be a result of the redundancy in speech signals. Because of this, the non-linearity of the auditory system is considered.

A brief overview of the auditory system is useful for understanding the data. The outer ear consists of the pinna and the ear canal; together they gather sound and provide directionality cues. The middle ear serves as an impedance transformer to couple sounds from the air to the liquid in the inner ear. The inner ear, or cochlea, amplifies the sounds and divides it up in to different frequency bands. Different frequencies then stimulate different nerve fibers which head to the brain. Additional processing takes places in the brain at the auditory cortex and in the cochlear nucleus.

To quantify some of these effects, two important types of data are examined here. In the cochlea, the basilar membrane moves up and down with the sound wave. The movements of this membrane make up one large set of the data. The other set of data is taken from the auditory nerve. The nerves that carry signals from the cochlea to the brain can be recorded. Data from the two sources is now believed to show agreement [32].

A: Total Harmonic Distortion

Prof. Dallos measured the harmonics in the auditory nerve fibers. The frequency of excitation was swept to find the peak response or "best frequency." Then the output of the hair cells was measured. Because the signal on the nerve fiber is made up of spikes, a large number of samples are taken and averaged together to form a composite signal. Then the Fourier analysis of the composite signal is taken to measure the harmonic content. At 30dB SPL, a sound level corresponding to a whisper, the harmonics are all between -40dB and -50dB with respect to the fundamental. This roughly 1% distortion level corresponds to the values used to define the maximum linear range in many analog filters. But, for sound levels corresponding to normal

speech, 50dB to 70dB, the third harmonic peak is near -20dB. Ten percent distortion is far from linear. [33].
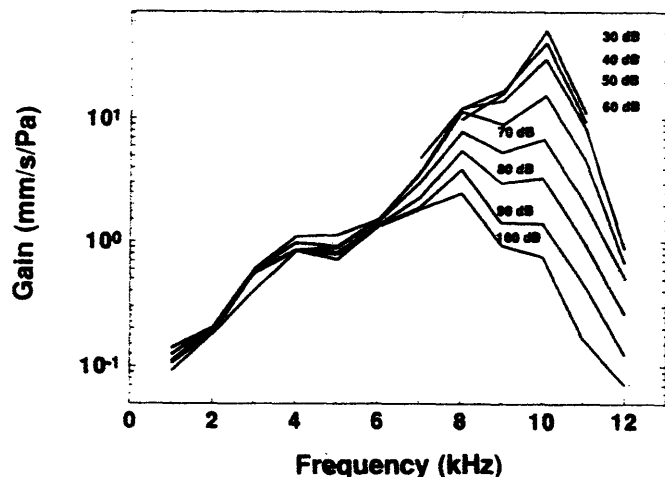
FIGURE 3.4. Basilar membrane velocity scaled by input sound pressure for a number of sound levels. Vertical overlap of the curves indicates linearity of the response. At lower frequencies. the curves align within approximately a factor of 2. and are characterized as being essentially linear. By contrast. in the region of the characteristic frequency (CF), the response is remarkably nonlinear. (Data from Ruggero. Rich. and Recio 1992.)

Figure 1.19: Basilar Membrane Non-Linearity, used with permission of Ruggero.

B: Other Nonlinear Effects

In addition to harmonic distortion, the auditory system demonstrates other behaviors that are non-linear. Figure 1.19 demonstrates the changing frequency response with amplitude. It shows the movement of the basilar membrane at one location as frequency and amplitude are varied. The Q decreases significantly and the center frequency decreases as the amplitude increases. Masking is the broad term for behavior in the auditory system where the presence of one sound prevents the listener from hearing another sound. This includes the non-linear effect called suppression. These effects are similar to desensitivity and blocking in RF systems, but occur at amplitudes where the system normally operates.

Section 5. Companding Speech Processor

Some of these nonlinear effects can be implemented in analog filters. Turicchia has proposed a filter system that captures two-tone suppression behavior [34]. Cochlear models have been built in the past [19], but because of the complex interactions in these systems, they can be

difficult to customize to applications such as speech processor front ends or cochlear implants. This simpler system shows promise for these applications.

B
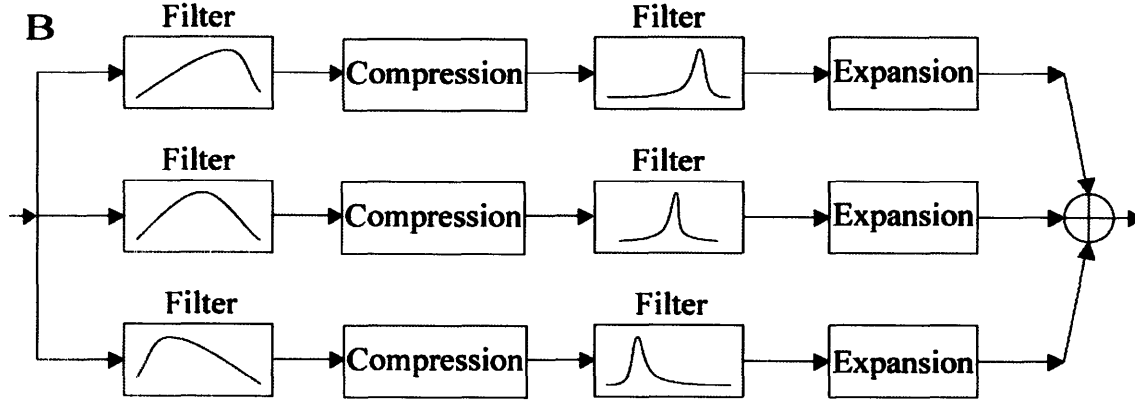
Figure 1.20: Companding Speech Processor

Turicchia's system is depicted in Fig. 1.20. A broad filter defines the frequencies that will influence the suppression. Then, non-linear compression is performed. A second filter selects the frequencies of interest. Then these frequencies are expanded.

This system can be well understood by examining a series of simple examples. For each of these examples, the system is set up to compress with the fifth root, $n_1 = 0.2$, and then expand to the fifth power, $n_2 = 1$. Initially, it is assumed that the frequencies being observed pass through both filters with no gain or attenuation. First, an input composed of a single sine wave of amplitude 1 is considered.

$$x_0 = 1\sin(\omega t) \tag{1.54}$$

The output of the compressor is the same as the input.

$$x_2 = x_0 = 1\sin(\omega t) \tag{1.55}$$

And the output of the expander is the same as the input as well.

$$x_4 = x_0 = 1\sin(\omega t) \tag{1.56}$$

A second possible input is a sine wave with an amplitude of 0.01.

$$x_0 = 0.01\sin(\omega t) \tag{1.57}$$

The compressor compresses the range of amplitudes by increasing this amplitude.

43

$$x_2 \approx 0.4\sin\left(\omega t\right) \tag{1.58}$$

Then the expander expands the range of amplitudes to recreate the original signal.

$$x_4 = x_0 = 0.01\sin\left(\omega t\right) \tag{1.59}$$

Both of these signals pass through the system with no amplitude change. The nonlinearities cancel for signals that are passed through both filters. Suppression occurs in more complicated signals where part of the signal is present in the compression stage, but removed before the expansion stage. A signal composed of the sum of the two earlier signals may be considered.

$$x_0 = 1\sin\left(\omega_1 t\right) + 0.01\sin\left(\omega_2 t\right) \tag{1.60}$$

The envelope of this signal is approximately 1, so the compression gain is one. And similar to Eq. 1.55,

$$x_2 = x_0 = 1\sin\left(\omega_1 t\right) + 0.01\sin\left(\omega_2 t\right) \tag{1.61}$$

If the second filter, G, removes the component at ω_1, then

$$x_3 = 0.01\sin\left(\omega_2 t\right) \tag{1.62}$$

The expansion step further decreases this amplitude.

$$x_4 = 10^{-10}\sin\left(\omega_2 t\right) \tag{1.63}$$

As in the biological cochlea, the gain control mechanism responds to large frequency components at nearby frequencies. When these are later filtered out, the small signal is smaller than it would have been had it been processed in the absence of the large signal.

This companding system is currently being tested by Turicchia in conjunction with collaborators for use as a preprocessor for cochlear implants and as a front end for a speech recognition system. To test the system for cochlear implant patients, he is working with Dr. Oxenham of the Auditory Perception and Cognition group at MIT's Research Laboratory of Electronics. Their experiment consists of sample sounds files that were recorded in a quiet environment. Noise is added to these samples. Then the companding processor is used to construct stimulation waveforms for each of the cochlear implant channels. In order to test the system with normal hearing subjects, a cochlear implant simulator is used to reconstruct audio samples. Subjects are presented the samples and given a forced choice.

44

Two sets of samples were used in the preliminary study: synthetic vowels and natural consonants in separate trials. In both cases, statistically significant improvements were demonstrated for low signal to noise levels as shown in Fig. 1.21. Future experiments in the study will examine performance on full sentence recognition with a variety of noise sources. To explore applications for computer speech recognition, Turicchia is working with researchers at the Mitsubishi Electronic Research Laboratory.
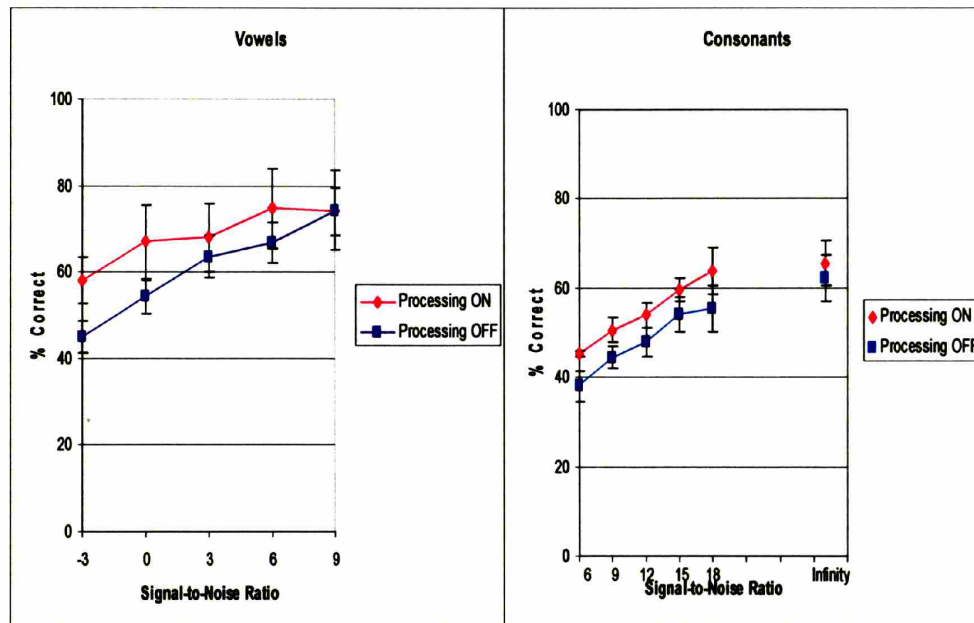
Figure 1.21: Companding Study Preliminary Results

Section 6. New Work

This thesis presents a series of work on analyzing and building nonlinear micropower filters. First, a capacitive-attenuation scheme is presented for increasing the linear range of filters. Then the nonlinear behavior of jump resonance is analyzed. Finally, a companding system is demonstrated that uses these filters.

A: Capacitive-Attenuation Filter

The initial gm-C design is part of a larger design for a cochlear implant speech processor. This system provides real world constraints on the filter's specifications: The layout must be compact to fit in the system chip; mismatch between the sixteen filters on each chip and

45

between chips is a concern; and, the programming mechanism will be tested. The capacitive-attenuation technique is used to match the linear range of the surrounding circuitry.

B: Jump Resonance and Automatic Q

The linear filters work well at low Q, but it becomes difficult to maintain linear range at high Q. This limitation is clearest in the case of lowpass or highpass filters. For these filters, the peak gain approaches Q for high Q filters. This gain means that smaller input signals are amplified to the point of distortion at the output. Bandpass filters generally exhibit a similar input linear range limiting effect since they typically encompass a highpass or lowpass filter in the feedback loop.

When the peak signals are high, the systems demonstrate distortion and a behavior referred to as jump resonance. This is a hysteresis in the frequency response near the peak. I present an analysis of this behavior and an automatic Q control circuit for avoiding jump resonance.

Automatic Gain Controls (AGCs) are a common solution to limited dynamic range. A variable gain amplifier is programmed based on the signal amplitude. AGCs form a special type of non-linearity. Because the gain is constant for a given signal amplitude, subject to the limitations of the amplitude detector, the system can be treated as a linear block for that amplitude.

Automatic Q Controls (AQCs) work the same way. Given a small input signal, the system acts like a linear filter with a high Q. As the input amplitude increases, the Q decreases. This ensures that harmonics are limited in the circuit.

C: Companding Speech Processor

I finally present a companding speech processor based on the algorithms being test by Turicchia. It uses the same block description presented in Fig. 1.20. It has sixteen channels, each with two filters, compression, and expansion blocks.

References

[1] W. K. Chen, *Passive and Active Filters: Theory and Implementations*. New York: John Wiley & Sons, 1986.

[2] A. S. Sedra and K. C. Smith *Microelectronic Circuits*. New York: Oxford University Press, 1997.

[3] T. Chen, "The Past, Present, and Future of Speech Processing". *IEEE Sig. Process. Mag.*, May 1998, pp. 24-48

[4] T. Painter and A. Spanias, "Perceptual Coding of Digital Audio". *Proceedings of the IEEE*. Vol. 88, No. 4, April 2000, pp. 451-513

[5] F. Baumgarte, "Improved Audio Coding Using a Psychoacoustic Model Based on a Cochlear Filter Bank". *IEEE Tran. On Speech and Audio Process.*, Vol. 10 No. 7, October 2002, pp. 495-503

[6] P. M. Clarkson and S. Bahgat, "Real-Time Speech Enhancement System Using Envelope Expansion Technique". *Electronics Letters* Vol. 25 No. 17, August 1989, pp. 1186-1888

[7] M. S. Ahmed, "Comparison of Noisy Speech Enhancement Algorithms in Terms of LPC Perturbation". *IEEE Trans. On Acoustics, Speech, and Signal Processing*, Vol. 37 No.1, Jan. 1989, pp.121-125

[8] S. Jeong and M. Hahn, "Speech Quality and Recognition Rate Improvement n Car Noise Environments". *Electronics* Letters, Vol. 37 No. 12, June 7th, 2001, pp.800-802

[9] F. A. Spellman, "The Past, Present, and Future of Cochlear Prostheses". *IEEE Eng. In Med. And Bio.* May/June 1999, pp. 27-33

[10] P.C. Loizou, "Mimicking the Human Ear". *IEEE Sig. Process. Mag.* Sept. 1998, pp. 101-130

[11] S.U. Ay, F.-G. Zeng, and B. J. Sheu, "Hearing with Bionic Ears". *IEEE Circuits & Devices*, May 1997, pp. 18-23

47

[12]E. Sanchez-Sinencio and J. Silva-Martinez, "CMOS Transconductance Amplifiers, Architectures and Active Filters: a Tutorial". *IEE Proc.-Circuits Devices Syst., Vol. 147, No. 1, Feb. 2000, pp. 3-12*

[13] P. Hasler, M. Kucic, and B. A. Minch, "A Transistor-Only Circuit Model of the Autozeroing Floating-Gate Amplifier". *Circuits and Systems, 1999. 42nd Midwest Symposium on,* Vol. 1, Aug. 1999, p 157-160

[14] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor,* New York, McGraw Hill, 1987

[15] C. Yoo, S.-W. Lee, and W. Kim, "A $\pm$1.5-V, 4-Mhz CMOS Continuous-Time Filter with a Single-Integrator Based Tuning". *Solid State Circuits, IEEE Journal of,* Vol. 33 No. 1, Jan. 1998, pp. 18-27

[16] A. Zeki, "Low-Voltage CMOS Triode Transconductor with Wide-Range and Linear Tunability". *Electronics Letters.* Vol. 35, No. 20, Sept. 1999, pp. 1685-1686

[17] U. Gatti, F. Maloberti, G. Palmisano, and G. Torelli, "CMOS Triode-Transistor Transconductor for high-frequency continuous-time filters". *Circuits, Devices and Systems, IEE Proceedings,* Vol. 141, No. 6, Dec. 1994, pp. 462-468

[18] A. Gharbiya and M. Syrzycki, "Highly Linear, Tunable, Pseudo Differential Transconductor Circuit for the Design of Gm-C Filters". *Proceedings of the 2002 IEEE Canadian Conference on Electrical & Computer Engineering,* 2002, pp. 521-525

[19] R. Sarpeshkar, R. F. Lyon, and C. Mead, "A Low-Power Wide-Linear-Range Transconductance Amplifier". *Analog Integrated Circuits and Signal Processing.* No. 13, 1997, pp. 123-151

[20] D. R. Frey, "Exponential State Space Filters: A Generic Current Mode Design Strategy". *IEEE Trans. On Circuits and Systems I.* Vol. 43 No. 1, Jan. 1996, pp. 34-42

[21] D. R. Frey, "On Log Domain Filtering for RF Applications". Bipolar/BiCMOS Circuits and Technology Meeting, 1995., Proceedings of the. Oct. 1995, pp. 12-18

[22] J. Mulder, W.A. Serdijn, A.C. van der Woerd, and A. H. M. van Roermund, "A Generalized Class of Dynamic Translinear Circuits". *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing.* Vol. 48 No. 5, May 2001, pp. 501-504

[23] E. M. Drakakis, A. J. Payne, and C. Toumazou, "'Log-domain state-space': a systematic transistor-level approach for log-domain filtering". IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing. Vol 46 No. 3, March 1999, pp. 290-305

[24] E. M. Drakakis and A. J. Payne, "A Bernoulli Cell-Based Investigation of the Non-Linear Dynamics in Log-Domain Structures". *Analog Integrated Circuits and Signal Processing,* No. 22, 2000, pp. 21-39

[25] D. R. Frey and A. T. Tola, "A State-Space Formulation for Externally Linear Class AB Dynamical Circuits". *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing.* Vol. 46 No. 3, March 1999, pp. 306-314

[26] G. Groenewold, "The Design of High Dynamic Range Continuous-Time Integratable Bandpass Filters". *IEEE Transactions on Circuits and Systems,* Vol. 38 No. 8, Aug. 1991, pp. 838-852

[27] J. B. Hughes, K. W. Moulding, J. Richardson, J. Bennett, W. Redman-White, M. Bracey, and R. S. Soin, "Automated Design of Switched-Current Filters". *IEEE Journal of Solid State Circuits.* Vol. 31 No. 7, July 1996, pp. 898-907

[28] G. Palaskas and Y. Tsividis, "A 'Divide and Conquer' Technique for the Design of Wide Dynamic Range Continuous Time Filters". *The 2001 IEEE International Symposium on Circuits and Systems. ISCAS 2001.* Vol. 1, May 2001, pp. 252-255

[29] M. Helfenstein, A. Muralt, G. Fischer, P. Zbinden, D. Pfaff, F. Frey, and G. S. Moschytz, "SC and SI Filters in Baseband Applications: A Comparison". *1997 IEEE Symposium on Circuits and Systems.* June 1997, pp. 297-300.

[30] T. Georgantas, S. Bouras, D. Dervenis, ad Y. Papananos, "A Comparison Between Integrated Current and Voltage Mode Filters for Baseband Applications". *The 6th IEEE International Conference on Electronics, Circuits and Systems, 1999.* Vol. 1, Sept. 1999, pp. 485-488

[31] B. Razavi, *RF Microelectronics* Upper Saddle River, NJ: Prentice Hall PTR, 1998.

[32] J. O. Pickles, *An Introduction to the Physiology of Hearing* Boston: Academic Press, 1988.

[33] B. C. J. Moore and R. D. Patterson, *Auditory Frequency Selectivity.* New York: Plenum Press 1986.

[34] L. Turicchia and R. Sarpeshkar, "A Bio-Inspired Companding Strategy For Spectral Enhancement," *IEEE Trans. On Biomed. Eng.,* Vol. 13, No. 2 March 2005 pp. 243-253

Chapter 2 : Capacitive Attenuation Filter[*]

Bionic ears (BEs) or Cochlear Implants have been implanted in more than 20,000 people [1]. They mimic the function of the ear in stimulating neurons in the cochlea in response to sound. Figure 2.1 shows an overview of a common signal-processing chain: sound is first sensed by a microphone. Pre-emphasis and gain control are then performed on the input. Bandpass filters (BPFs) then divide the sound into different frequency bands. The envelope of each channel is detected. The dynamic range of each channel is then compressed to fit into the patient's dynamic range. The signals from each channel are then modulated and sent to the electrodes to stimulate the remaining neurons in the patient [2].

Bionic Ear Overview
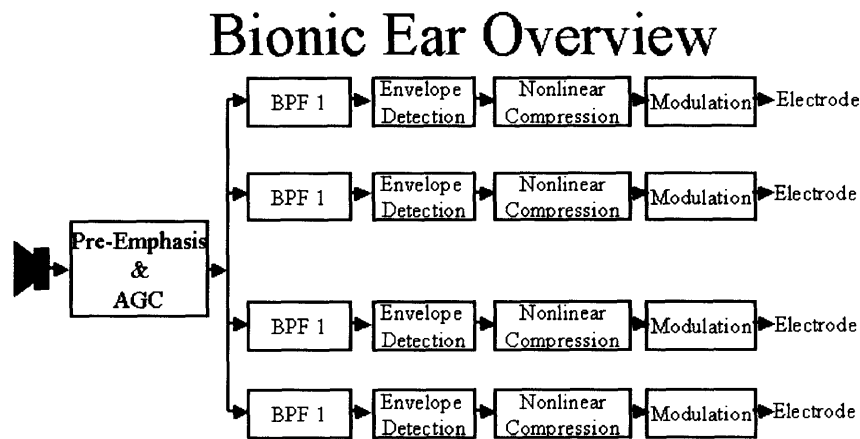
Figure 2.1: Bionic Ear Overview

Current systems use a DSP-based processor that may be worn as a pack on the belt or as a Behind-The-Ear unit. The challenge is to now move to designs that can be fully implanted. Reducing the power of the BE is one of the keys to moving to a fully implanted system, and provides an important motivation for this work.

We would like to implement bandpass filters with microwatt and submicrowatt power consumption rather than the hundreds of microwatts that are typical of current DSP implementations.

Bionic Ear Filter Requirements

While the signal processing in current BEs is performed in DSPs, analog computation has been proposed as a mechanism for decreasing the power used in these steps [3],[4],[5]. The BE application offers a number of constraints on the design of bandpass filters. It is battery powered and required to run off a low voltage this design is optimized for 2.8 Volts. The filter must be tunable over most of the audio range, from 100Hz to 10kHz. It should have a dynamic range of at least 60dB. And it must minimize power while achieving these specifications.

Subthreshold Gm-C filters were proposed for this application in [3],[4] because of the wide tuning range and low power. The drawback of this choice is the small linear range of the transconductors, approximately 75mV for a simple OTA. This forces the surrounding circuitry to interface to the filters with small signals that are prone to noise and other effects. A combination of degeneration, attenuation, and non-linear term cancellation were proposed to increase the linear range in some of our prior work [6]. All of these techniques are aimed at decreasing the 3rd harmonic that forms due to the saturation of the input differential pair and which limits the linear range of a simple transconductor. The combined techniques meet the linear range requirements listed above, but have limitations. The linear range cannot be easily increased above the achieved value because the techniques do not scale. Because much of the increase in linear region comes from using the well as an input, the linear range is a function of the subthreshold body-effect parameter, κ, which is poorly regulated between processes and even between runs. As shown in Fig. 2.2, the 2nd harmonic becomes significant at amplitudes well below where the 3rd harmonic limits performance.
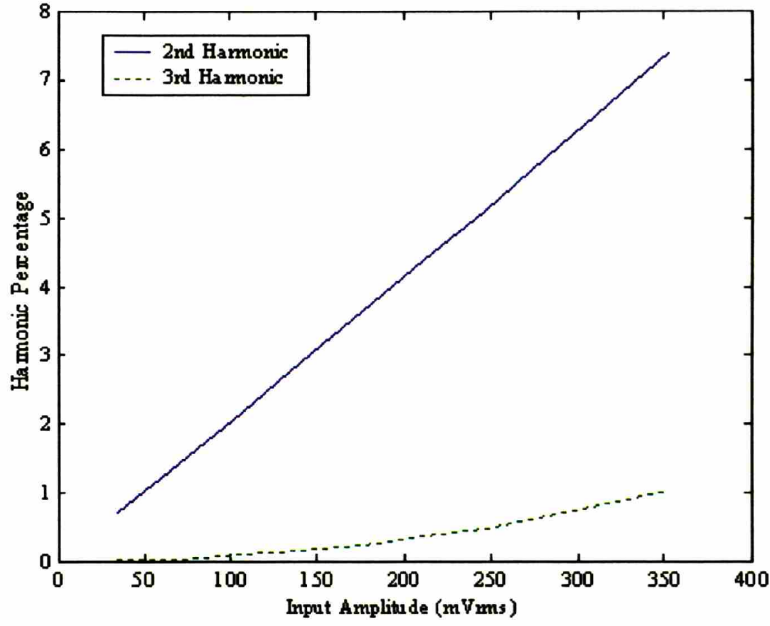
Figure 2.2: Harmonics from Bandpass Filter built using Wide Linear Range Transconductors [3]

This paper demonstrates bandpass filters that extend the linear range of transconductors using passive attenuators built with capacitors as proposed in [6]. This is similar to the technique of passive attenuation using resistors that is found in some discrete transconductors [7]. The passive attenuator ensures that the attenuator does not introduce harmonic distortion and allows for simple scaling of the attenuation ratio to adjust the linear range. However, care must be taken to set the D.C. values of floating nodes, as we will discuss below.

Basic Capacitive-Attenuation Filter Topology

The basic bandpass topology examined here is a cascade of first-order highpass and first-order lowpass filters based on RC primitives (Fig. 2.3). The low pole is proportional to the bias current of G_1, while the high pole is proportional to the bias current of G_2. The major limitation of this design is its small linear range. Signal amplitude is limited to the linear range of the transconductors.
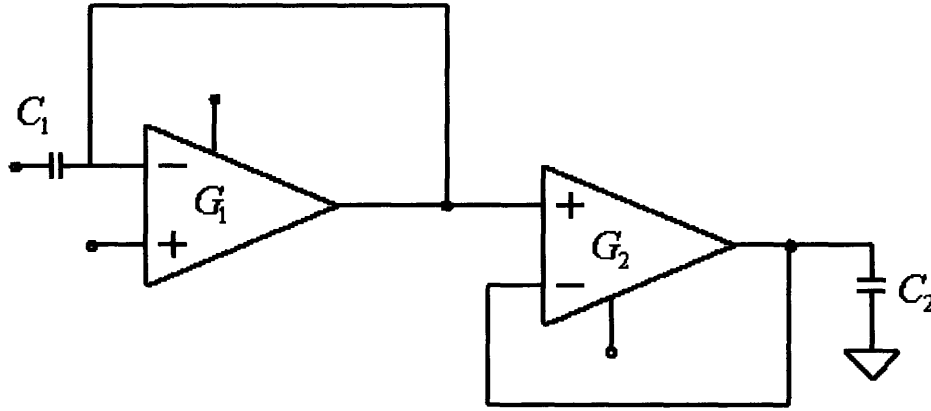
Figure 2.3: Basic Bandpass Topology

The linear range can be improved by attenuating the input and attenuating the output in feedback as shown in Fig. 2.4 because the linear range of the transconductors is limited by their input swing, rather than their output swing. In the highpass stage, the signal is attenuated by a factor of A+1, where A is the ratio of the attenuator capacitances. The full capacitance of $(A+1)C_1$ is then used for filtering with G_1. In the lowpass stage, a gain of A+1 is applied to signals in the passband. A capacitance C_3 is added in parallel with the attenuating capacitances to increase the filtering capacitance and lower the reverse transmission from the inverting node of G_2 to V_{out}. The transfer function of the circuit in Fig. 2.4 is

$$H(s) = \frac{sC_1(A+1)G_2}{\left[sC_1(1+A)+G_1\right]\left[G_2+s\left(C_3(A+1)+AC_2\right)\right]} \cdot \tag{2.1}$$
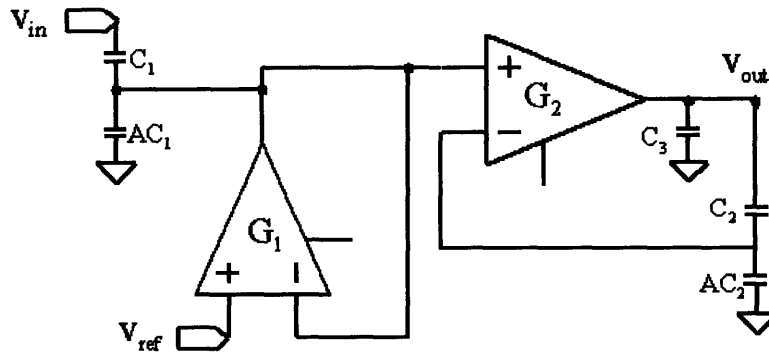
Figure 2.4: Capacitive Attenuation Bandpass Filter

54

While this design has the proper transfer function, it will not function well in practice because the inverting input to G_2 does not have a well-defined DC value. The transconductor has a gate of a MOSFET at its input and the other two components attached to the node are capacitors; similarly, the output of G_2 is poorly constrained. Adding a transconductor, G_3, that constructs a weak low-frequency path between the two nodes constrains both DC values, as shown in Fig. 2.5. The transconductor G_3 operates on full-scale signals unlike G_1 and G_2 that operate on attenuated signals. Hence, for large signals, G_3 will saturate and introduce distortion. The low bias current of G_3 makes that distortion negligibly small.
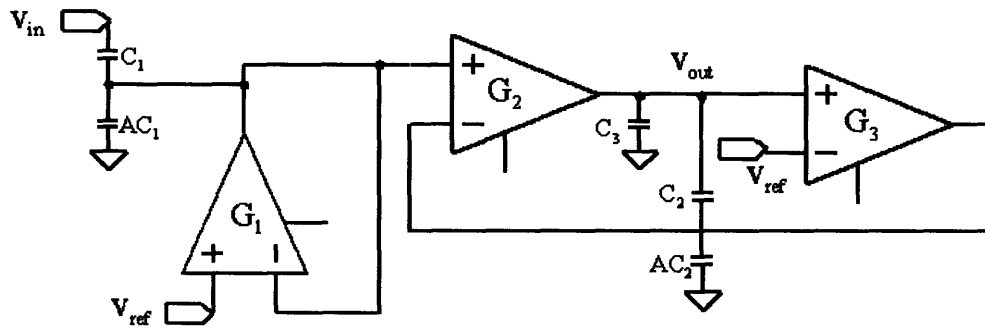
Figure 2.5: Capacitive-Attenuation Filter With Offset Adaptation
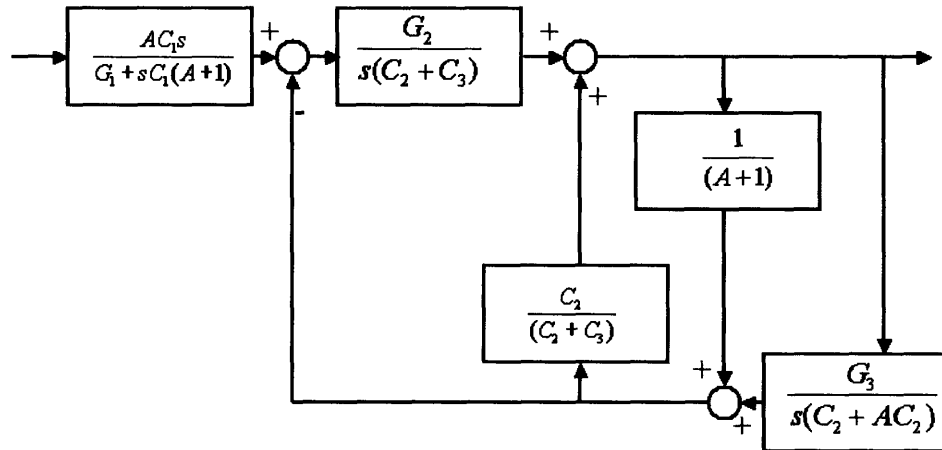
Figure 2.6: Block Diagram Of Capacitive Attenuation Filter With Offset Adaptation

The block diagram in Fig. 2.6 for the circuit filter in Fig. 2.5 shows another area for concern with this design: A capacitive path from the output of G_3 to its non-inverting input provides positive feedback. If G_3 were configured as a resistor this path would be guaranteed to have a

gain of less than one, ensuring stability. With the current topology, the possibility for instability does exist, but it is avoided because of the low bias current for G_3. This possibility of instability is also seen in the transfer function. The Routh criterion states that all coefficients in a second-order polynomial must have the same sign to ensure all the poles are in the left half plane [7]. The s term of the second portion of the transfer function becomes negative if $G_3 > G_2$.

$$H(s) = \left(\frac{sC_1}{G_1 + sC_1(A+1)} \right) \left(\frac{sG_2C_2(A+1)}{G_2G_3 + sC_2(G_2 - G_3) + s^2((A+1)C_2C_3 + AC_2^2)} \right) \quad (2.2)$$

Theoretical Noise Analysis of Basic Bandpass Topology

In order to size the devices, an attenuation ratio is first chosen based on the desired linear range of the system. Capacitor values are then selected to set the thermal kT/C noise at the required level for the desired dynamic range. The required current levels for the desired pole locations dictate the transistor sizing necessary to maintain fully subthreshold operation and a dominance of thermal over $1/f$ noise [6]. A signal swing of nearly $\pm 1V$ can be achieved with an attenuation ratio of 12; given a transconductance amplifier linear range of 75mV; then, A=11.

Noise analysis begins with a calculation of the input-referred noise from each transconductor. Simple transconductors have four transistors that contribute noise to the output. In equilibrium each transistor carries half of the bias current, yielding an output current noise of

$$i^2_{noise,out} = 4\left(2q\left(\frac{I_b}{2}\right)\right) = 4qI_b \quad (2.3)$$

This current noise can be referred to the input as voltage noise by dividing by the transconductance squared.

$$v^2_{noise,in} = \frac{4qI_b}{G^2} = \frac{4qV_L^2}{I_b}, \quad (2.4)$$

where $V_L \approx 75mV$ is the linear range of the transconductor.

56

The total noise at the output is found by applying the input-referred noise sources from each of the transconductors to the circuit and summing the effect of each noise source on the output. This can be visualized in the block diagram in Fig. 2.7 or using the following equations.
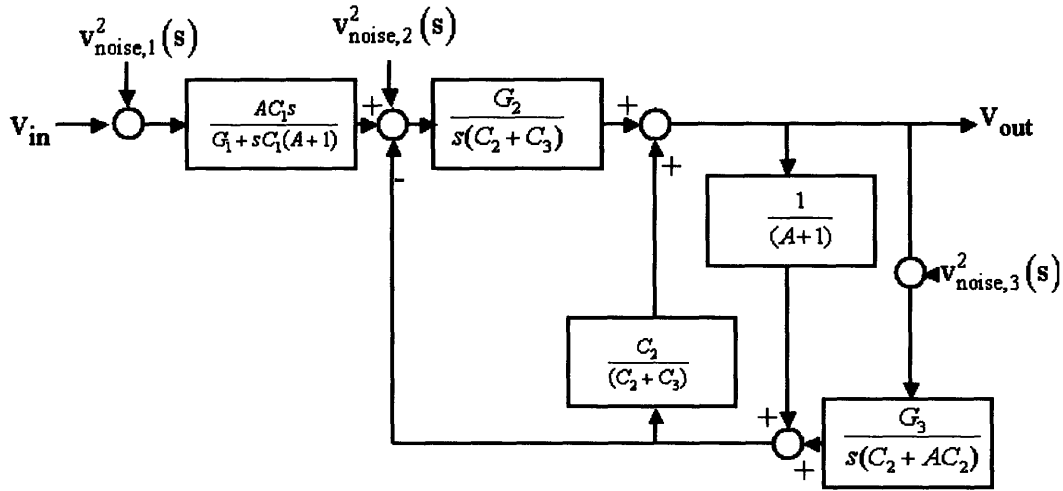
Figure 2.7: Block Diagram of Single Capacitive-Attenuation Filter With Noise Sources

$$v_{noise,tot}^2 (s) = v_{noise,1}^2 (s) + v_{noise,2}^2 (s) + v_{noise,3}^2 (s) \tag{2.5}$$

$$v_{noise,1}^2 (s) = \left(\frac{4qV_L^2}{I_{b,1}} \right) \left(\frac{G_1}{G_1 + s(A+1)C_1} \right)^2 \text{Lowpass}^2 (s) \tag{2.6}$$

$$v_{noise,2}^2 (s) = \left(\frac{4qV_L^2}{I_{b,1}} \right) \text{Lowpass}^2 (s) \tag{2.7}$$

where,

$$\text{Lowpass}(s) = \frac{sG_2 C_2 (A+1)}{G_2 G_3 + sC_2 (G_2 - G_3) + s^2 ((A+1)C_2 C_3 + AC_2^2)} \tag{2.8}$$

$$v_{noise,3}^2 (s) = \left(\frac{4qV_L^2}{I_{b,3}} \right) \left(\frac{G_3 (sC_3 - G_2)}{G_2 G_3 + sC_2 (G_2 - G_3) + s^2 ((A+1)C_2 C_3 + AC_2^2)} \right) \tag{2.9}$$

For hand calculations of the noise, we will neglect the effect of G_3. The noise from G_1 is the same as the noise from a lowpass filter with capacitance equal to $(A+1)C_1$. Since G_1 is setting the lower pole of the transfer function, its noise falls in the passband of the filter. The noise at the output from G_1 is then calculated by techniques similar to those described in [6]

$$v_{noise,1}^2 \approx \frac{2kT(A+1)^2}{\kappa(A+1)C_1} = \frac{2kT(A+1)}{\kappa C_1} \tag{2.10}$$

Similarly, if G_3 is assumed to be small the noise from G_2 takes the form

$$v_{noise,2}^2(s) = \left(\frac{4qV_L^2}{I_{b,2}}\right)\left(\frac{G_2(A+1)}{G_2 + s((A+1)C_3 + AC_2)}\right)^2 \tag{2.11}$$

This corresponds to a total integrated noise at the output from G_2 of

$$v_{noise,2}^2 \approx \frac{2kT(A+1)^2}{\kappa\left[(A+1)C_3 + AC_2\right]} \approx \frac{2kT(A+1)}{\kappa(C_3 + C_2)} \tag{2.12}$$

These hand calculations suggest that C_1 and C_2+C_3 should be around 6pF for a noise floor of 200uVrms at the output, which is necessary for a dynamic range of 70dB with a maximum signal of 2Vpp. More detailed computations in Matlab suggest that $C_1=C_2=5pF$ and $C_3=3pF$ are sufficient.

Transistor Sizing

This filter was designed around subthreshold transconductors because they offer a low saturation voltage, wide tuning range, and are typically thermal noise dominated [6]. To ensure that the transconductor is subthreshold, the devices must be scaled to accommodate relatively large subthreshold currents.

The pole locations are approximately

$$p_1 = \frac{G_1}{(A+1)C_1} = \frac{I_{b,1}}{(A+1)C_1V_1}, \tag{2.13}$$

58

$$p_2 = \frac{G_2}{(A+1)C_3 + AC_2} = \frac{I_{b,2}}{V_1\left[(A+1)C_3 + AC_2\right]}$$

<div align="right">(2.14)</div>

The maximum bias currents that will be required are at the high end of the audio frequency range where the poles are set at 5kHz and 10kHz. Those currents are computed to be 125nA and 430nA for G_1 and G_2. Those currents are large for subthreshold design. Accordingly, the first chip was built with 150um/3um transistors in the G_1 and G_2 transconductors. The transistors in G_3 did not need to be large since they carried a small bias current.

Experimental Results

A chip with this capacitive attenuation filter was fabricated on AMI's 1.5um process through MOSIS. Figure 2.8 shows that the filter can realize the ideal transfer function. The added pole and zero associated with G_3 are not visible. If the current in G_3 is turned up and the filter is set to lower frequencies by turning down the bias currents for G_1 and G_2, the influence of G_3 can be observed as in Fig. 2.9. We obtain a dynamic range of 68dB with 2uW power consumption for a 5kHz-10kHz filter. Table 1 reveals further details. The experimental dynamic range was approximately 2dB less than that predicted by theory.

Parameter	Value
Power (100Hz to 200Hz)	42nW
Power (5kHz to 10kHz)	2.0uW
Noise	323uVrms
Amplitude at 5% distortion	813mVrms
Dynamic Range	68.0dB

Table 1: Experimental Results For A Single Capacitive Attenuation Filter
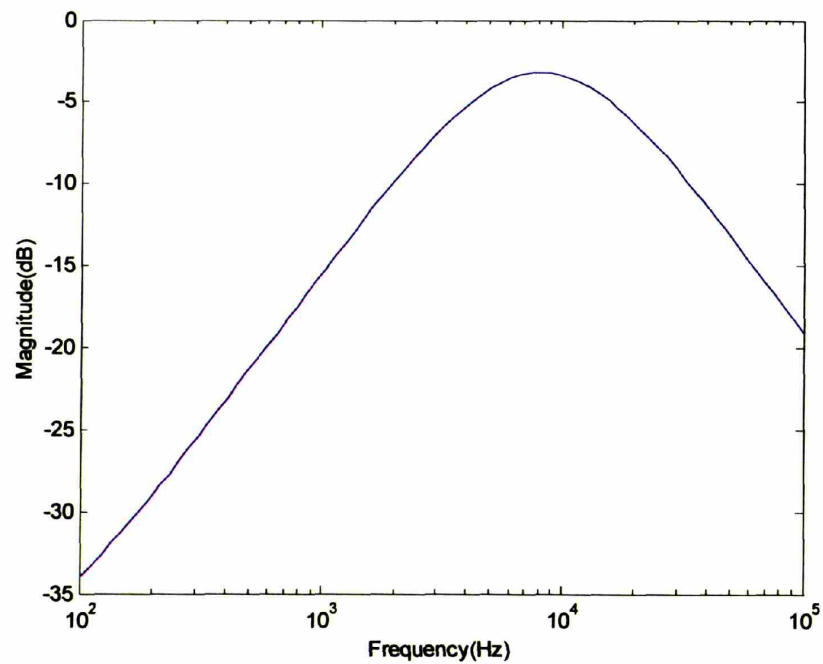
Figure 2.8: Experimental Transfer Function of Single Capacitive-Attenuator Filter
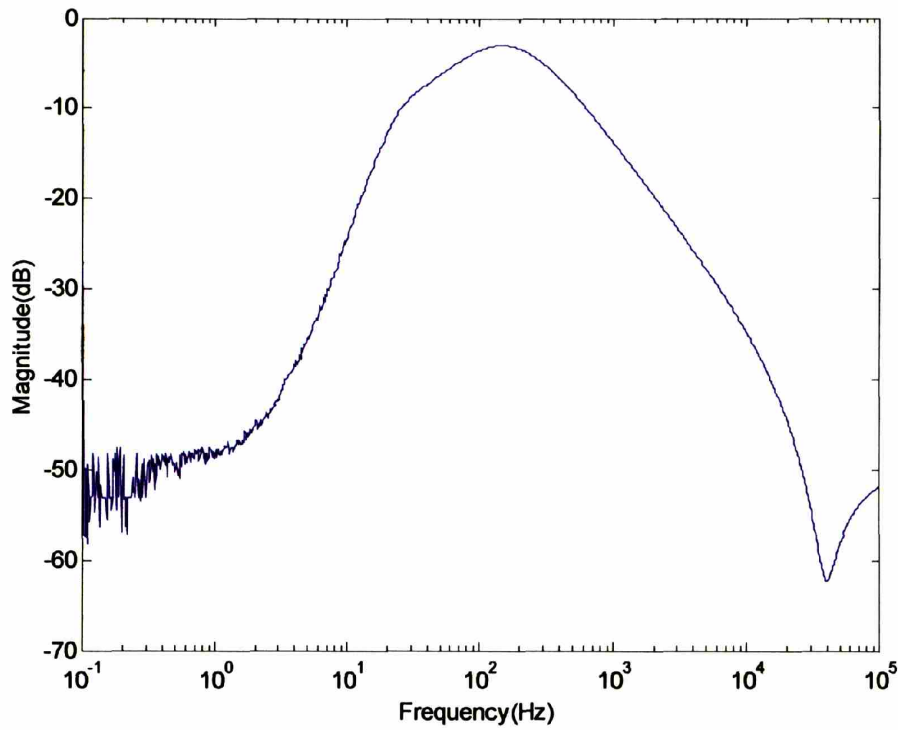
Figure 2.9: Experimental Demonstration of Offset Adaptation Effect on Transfer Function

The noise spectrum of the output node is plotted on top of simulation results from SPICE in Fig. 2.10. The figure demonstrates that the measured noise was accounted for by a fit to Equations (5) through (9). The initial low-frequency portion of the spectrum is due to Equation (9), while the later high frequency portion is due to Equations (6) through (8). Gratifyingly, the total integrated output noise was found to be 200uVrms which is in good agreement with our approximate prior hand calculations.
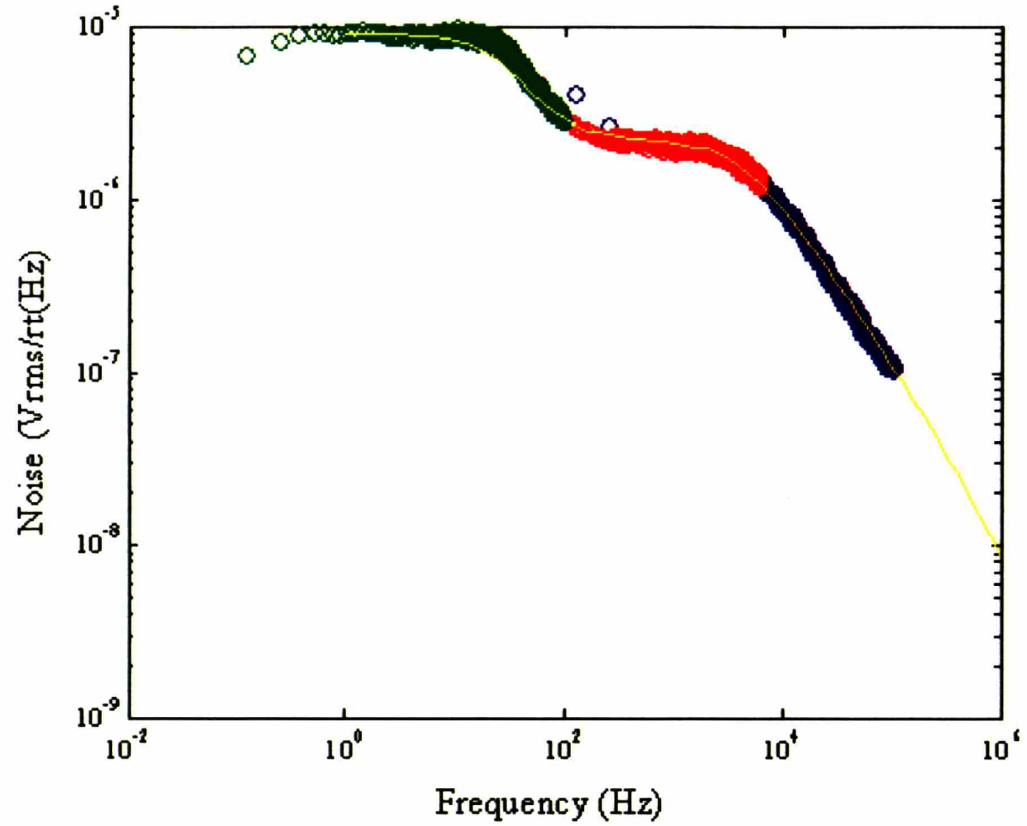
Figure 2.10: Measured and Simulated Noise for Single Capacitive Attenuation Filter

The linear range of the filter is defined to be the amplitude of the input signal at which the root-mean-square amplitude of the harmonics is 5% of the amplitude of the output fundamental signal. Figure 2.11 shows the percentage growth of second and third harmonics with input signal amplitude in rms. This figure demonstrates an improvement in the levels of second harmonic compared to the wide linear range transconductor data shown in Fig. 2.2. The second harmonic still dominates the third harmonic over the operating region, which suggests that a differential topology may be helpful. We now describe experimental results from a differential topology.
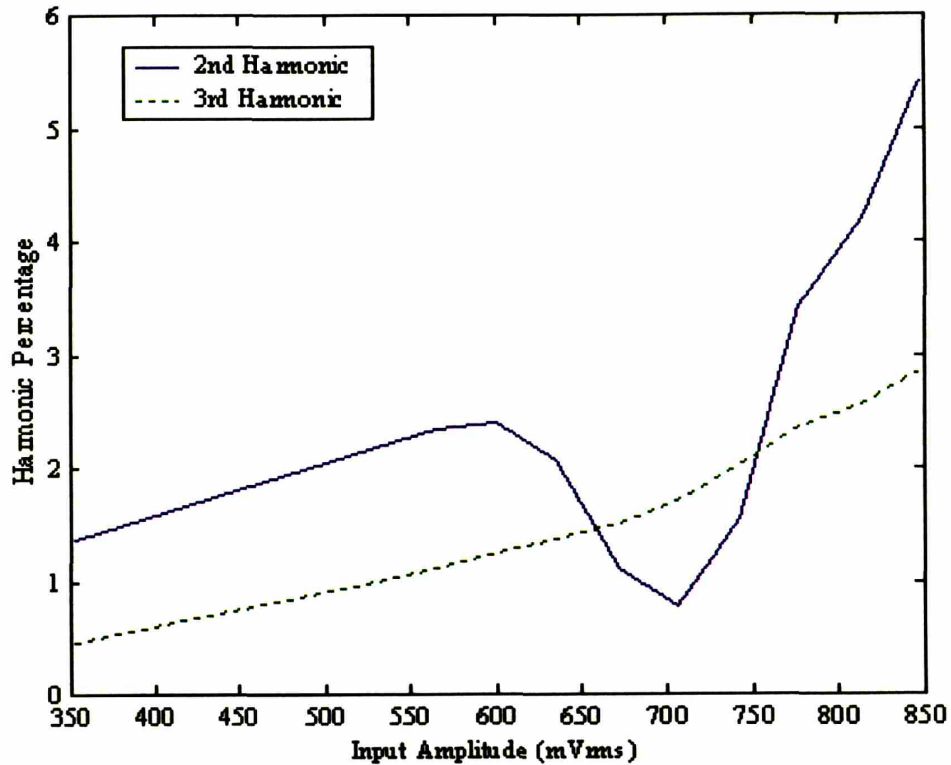
Figure 2.11: 2nd and 3rd Harmonics for Single Capacitive Attenuation Filter

Differential Topology

By processing a signal and its inverse through identical channels, differential topologies double the linear range of the composite signal and help cancel out common effects such as power supply noise and even harmonics. The primary motivation with our differential filter is to attenuate the second harmonics. Often when single-ended designs are converted to differential topologies, the two channels are combined into a common differential path using differential components and an additional common-mode path is created with lower performance requirements.

That approach was considered here, but a topology with two fully independent channels was pursued: The power in our channels is mostly being used to set the operating bandwidth and noise levels. The potential savings in overhead of a combined differential approach are minimal. Thus, our approach of two independent channels allows operation of half of the

63

system as a single-ended design if one of the channels fails. In an implanted system, such flexibility may be useful.

The drawback of the two-channel approach is that care must be taken to ensure that the two channels will match. In designs that use differential transconductors this only requires matching of pairs of transistors. With two independent channels, each component must be matched to its twin in the other channel. Figure 2.12 shows the layout of such a stage with one of the channels shaded. The capacitors are interspersed and the transconductors are intertwined at the transistor level. Figure 2.13 shows two such intertwined transconductors
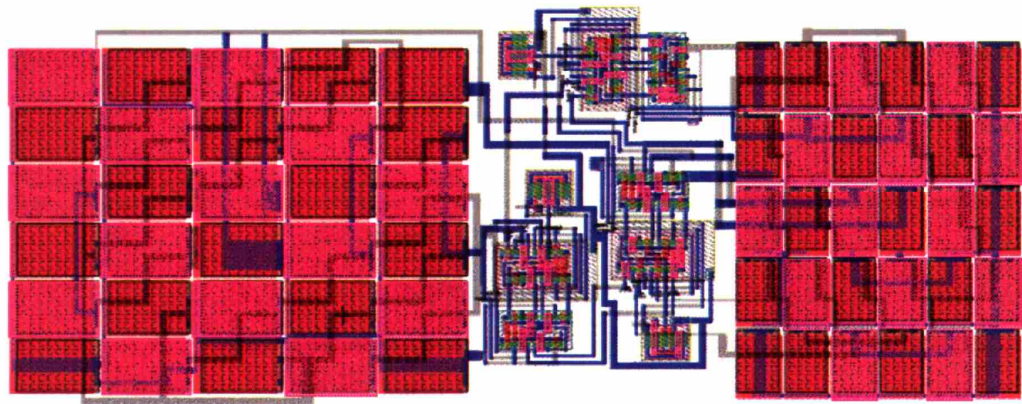
Figure 2.12: Two Interleaved Capacitive Attenuator Filters (One channel is shaded)
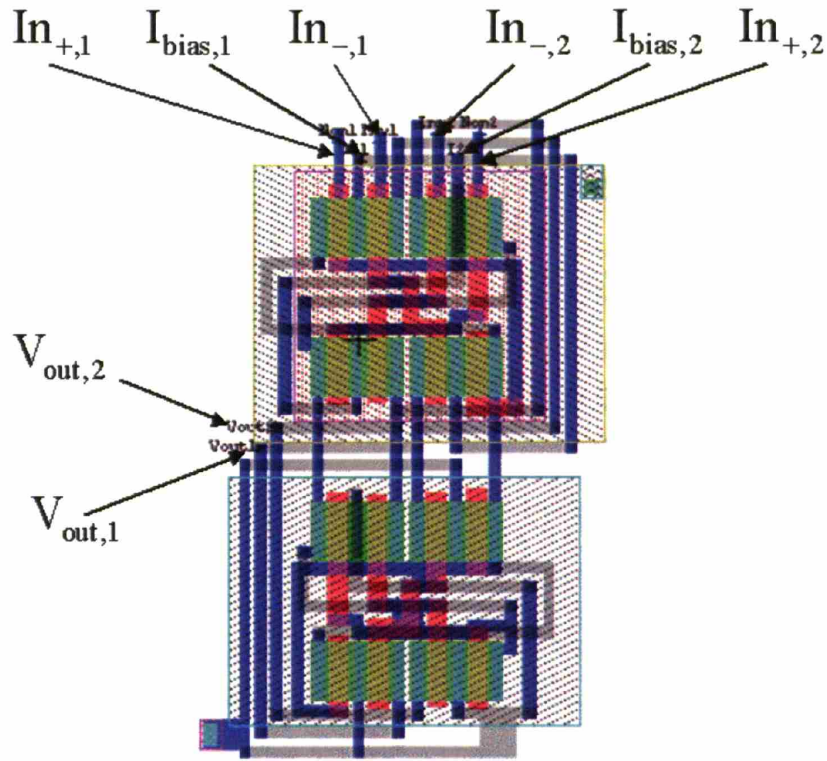
Figure 2.13: Two Intertwined Transconductors

In addition to the layout concerns, device sizing must be adjusted for the differential design. As noted earlier, the linear range of the system doubles with two channels even if there are no improvements from canceling the second harmonic. This improvement is simply because the composite signal is now the value of the noninverted channel minus the value of the inverted channel. When the noninverted channel is at its maximum, the inverted channel is at its minimum and the difference is twice the maximum of either channel. The noise from the two channels adds, but, because the noise is independent it adds as a sum of squares. Thus, the linear range doubles while the total noise only increases by a factor of $\sqrt{2}$. Since our single-channel system already demonstrated the required dynamic range, the capacitance of each channel may be decreased by a factor of two to lower the dynamic range of the differential scheme back to the single-channel value. Thus the total capacitance and total power of the differential system is the same as that for a single-ended system, but the linear range is greater because of the effect of second-harmonic attenuation. Indeed, out experiments reveal that the

65

differential filter has a dynamic range greater than 70dB at 3.36uW for a 5kHz-10kHz filter as shown in Table 2.

Parameter	Value (filter set 5kHz to 10kHz)
Power	3.36uW
Noise	535uVrms
Amplitude at 5% distortion	1.8Vrms
Dynamic Range	70.7dB

Table 2 : Experimental Results Of A Differential Capacitive Attenuation Filter

Each of the two differential channels performs similarly to the single channel presented in section III. The increase in linear range is a result of second harmonic attenuation. Figure 2.14 shows the expanded linear range and the fact that the second harmonic is quite low through most of the operating amplitudes. The second harmonic is below the third harmonic over a good range of amplitudes in contrast with the single ended results of Fig. 2.11.
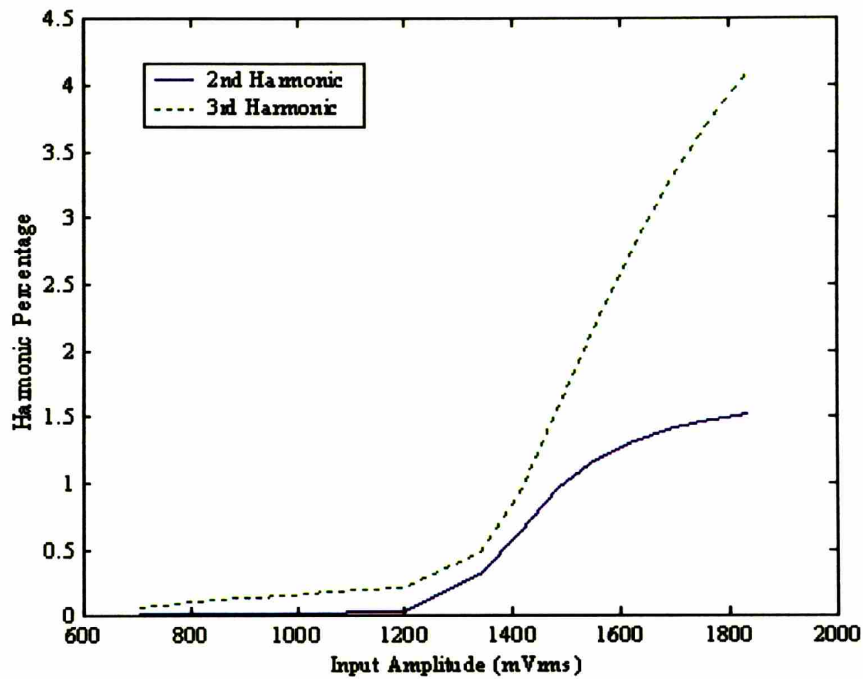
Figure 2.14: 2nd and 3rd Harmonics As A Function of Fundamental Amplitude For a First-Order Differential Capacitive-Attenuation Filter

Cascading Filters

Having demonstrated that a filter with first-order rolloff above and below the pass band can be built, the next step is to demonstrate that such filters may be cascaded to achieve higher-order rolloffs. The lowpass section does not load the highpass section in a single stage because the transconductor in the lowpass section also acts as a buffer. But, to prevent the highpass section of a subsequent stage from loading the lowpass section of a previous stage, an explicit buffer is necessary.
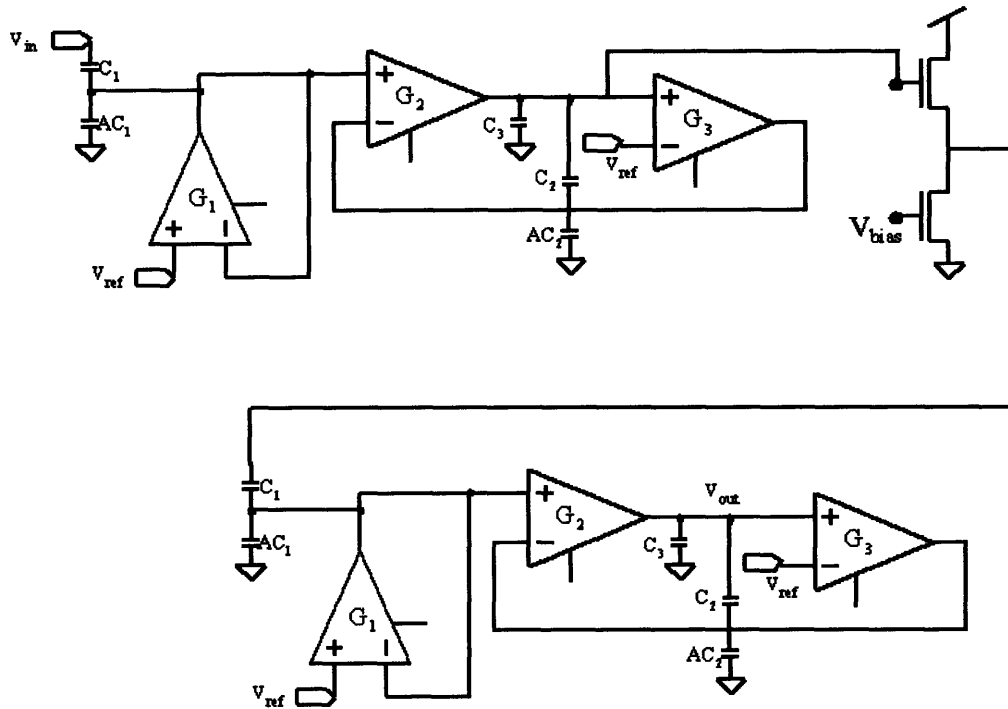
Figure 2.15: Two Capacitive-Attenuation Filters With An Intermediate Source Follower Buffer

This buffer is implemented with a simple source follower as depicted in Fig. 2.15. To ensure that the output of the buffer can be treated as a low impedance in our circuit, the buffer must have a much higher transconductance than the transconductors used in the filter stages. It is tempting to try and use a lower bias and simply include the buffer's output impedance in the filter design. The buffer is, however, operating on full size signals, and such an approach can significantly increase distortion. If, instead, the buffer is biased at the same current as G_2, the performance of the filter is not affected much because it is significantly higher bandwidth (greater by A+1 with respect to the preceding filter) allows the full-scale signal (greater by A+1 with respect to the preceding filter's input signal) to be passed through with little distortion.

While cascading two filters improves the roll off, it diminishes the dynamic range of the system by both increasing the noise and by diminishing the maximum undistorted signal at the input. The noise power could possibly increase by up to a factor of 2 because the number of devices contributing noise has doubled. However, some of the noise from the first stage is filtered by the second stage reducing the amount of the noise increase at the output. Detailed calculations

68

suggest that such filtering removes roughly half the noise power of the first filter for a one-octave filter.

The maximum undistorted signal at the output decreases because of the addition of harmonic content from the two stages and due to passband attenuation. The linear range decline due to harmonic addition in this design is minor because of the shape of the harmonics shown in Fig. 2.14: If the second stage does no harmonic filtering, harmonic addition in the two stages could possibly constrain the maximum signal to a point where the harmonics of a single stage are half the acceptable level. However, the harmonic content rises rapidly at the end of the linear range and this steep slope causes only a small loss in operating range when referred to the input. The passband attenuation inherent in our passive topologies poses a larger effect on octave filters, by lowering the passband signal by 3.5dB or equivalently by increasing the input referred noise by the same factor.

These effects combine to diminish the dynamic range of our second-order filter to slightly more than 60dB from the first-order result of slightly more than 70dB. The detailed results are shown in Table 3. Most BEs have a front-end gain control system that maps the 80dB dynamic range in the input to a dynamic range of 40dB to 60dB at the input of the BPFs. Thus, our second order filter's dynamic range is still within the specifications needed in BEs.

Experimental Results

Figure 2.16 shows experimental measurements of the second-order filter after its first stage and after its second stage. Theoretical fits to first and second-order filter transfer functions are also shown. The matching of the two stages appears to be excellent.
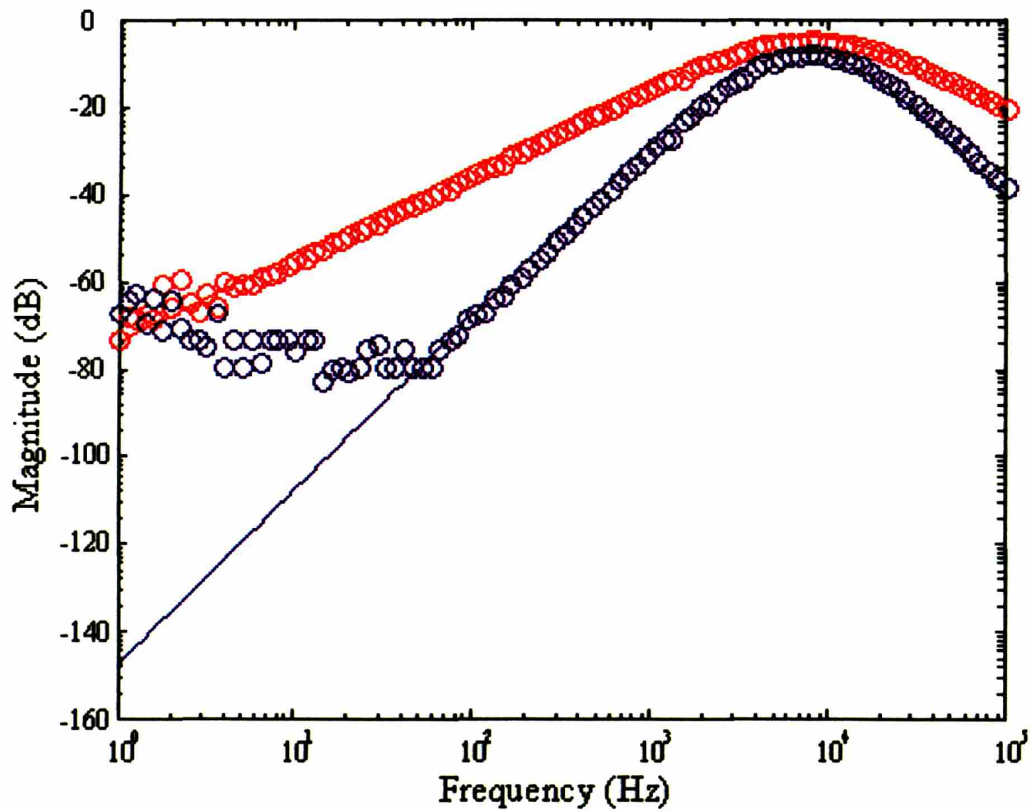
Figure 2.16: Fit Of First And Second Stage Outputs To Ideal Filter Transfer Functions

The noise increase with the addition of a second stage is shown in Fig. 2.17. The integrated output noise after the first stage and the buffer is 325uVrms, slightly more than was measured with a single stage alone. The integrated noise after the second stage is 422uVrms, just short of $\sqrt{2}$ times as much. The ability of the second stage to filter the noise from the first stage was less than predicted. It can be seen at high frequencies, however, that the two curves grow closer together, as the high frequency noise from the first stage is removed via filtering at the second stage.
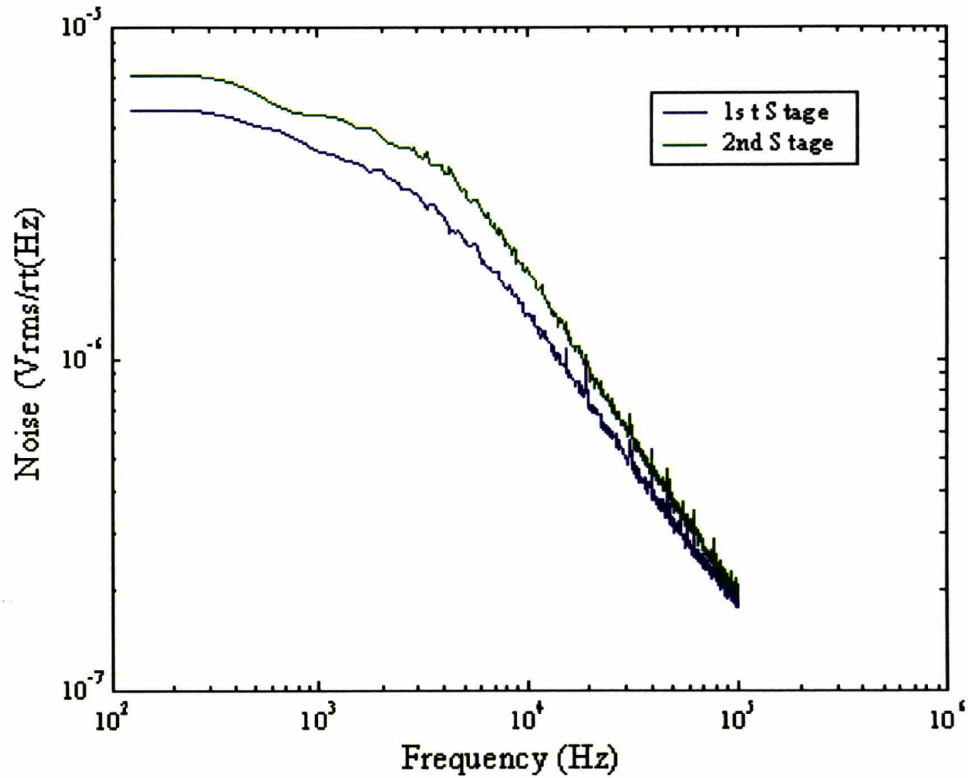
Figure 2.17: Noise In A Cascaded Differential Capacitive Attenuation Filter, After the First Stage And At The Output

Figure 2.18 shows the harmonic distortion plot for the second-order filter. The distortion does not reach five percent even for large signal amplitudes because the harmonics from the first stage are being filtered by the second stage to below 5% even when the first stage is being driven into saturation. This effect is more clearly shown by the decrease in gain with increasing input signal amplitude plotted in Fig. 2.19. Given these facts, the distortion point is a poor measure of the linearity of this filter. Instead, we choose the point where the gain begins to fall rapidly as the maximum undistorted input signal for the filter.

71
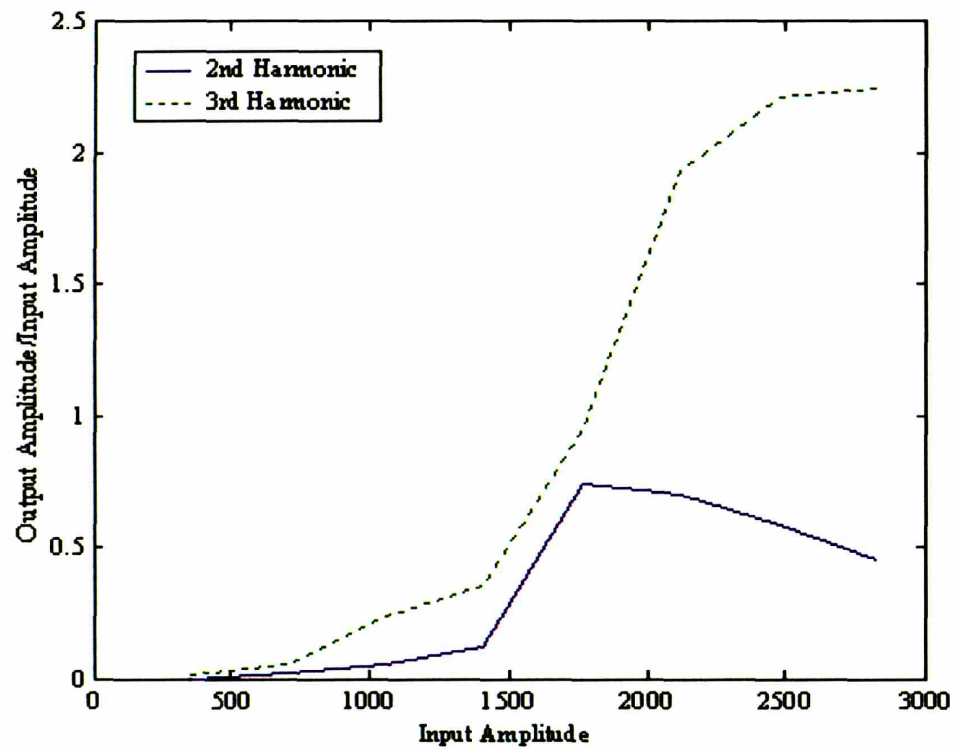
Figure 2.18: Harmonic Distortion Measurements for The Cascaded Differential Capacitive-Attenuation Filter
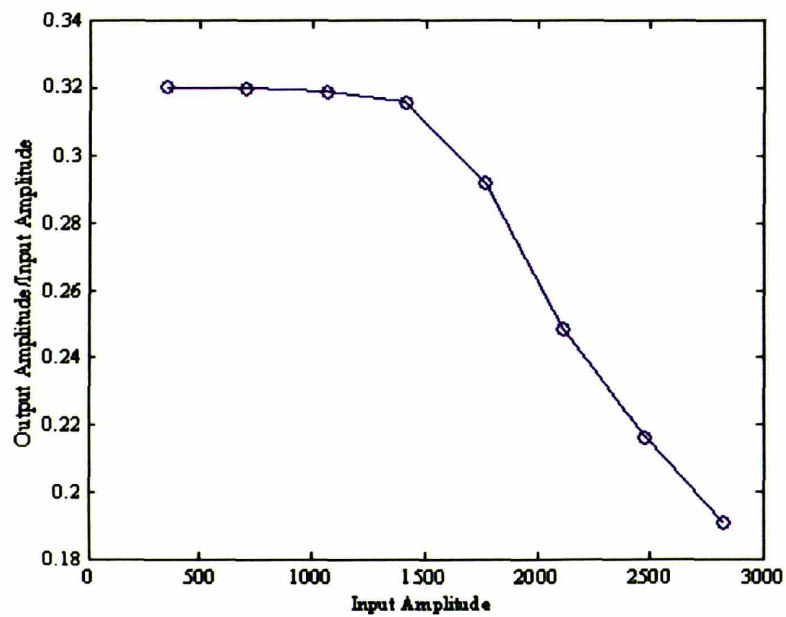
Figure 2.19: Gain vs. Input Amplitude for the Cascaded Differential Capacitive-Attenuation Filter

72

The ability to program the pole locations of the filter over a wide range was an important reason for picking the subthreshold Gm-C filter topologies. That programmability was tested using an on-chip 5-bit DAC to supply the bias currents for each pole. Figure 2.20 shows the frequency response of the second-order filter for each of the 32 values with the poles placed an octave apart.
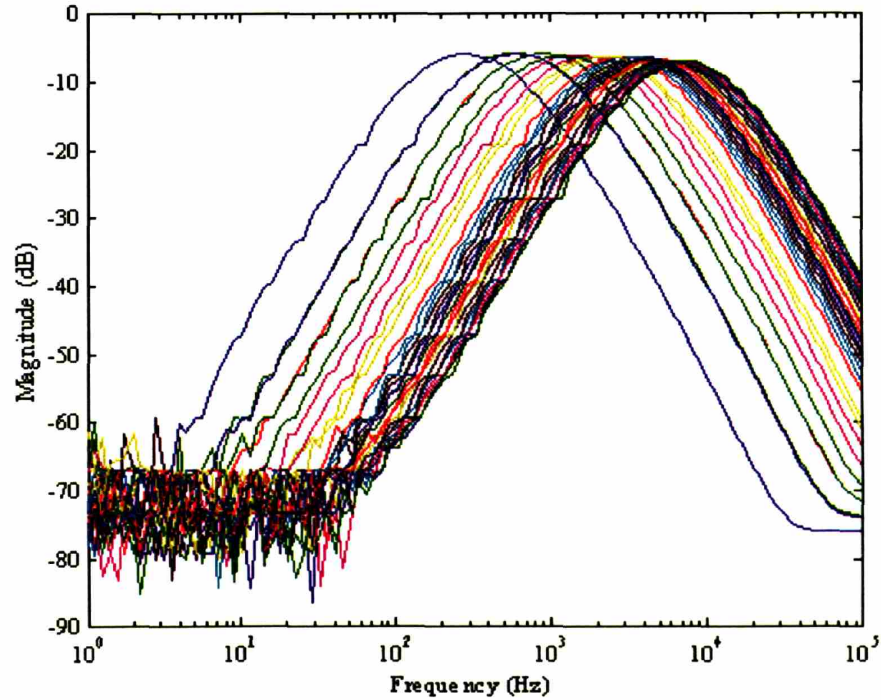
Figure 2.20: Demonstration Of Programmability Of The Cascaded Differential Capacitive-Attenuation Filter

Parameter	Value (filter set 100Hz to 200Hz)	Value (filter set 5kHz to 10kHz)
Power	230nW	6.36uW
Noise	1292uVrms	1318uVrms
Amplitude at 5% distortion	1414mVrms	1414mVrms
Dynamic Range	60.8dB	60.6dB

Table 3: Experimental Results Of A Cascaded Differential Capacitive Attenuation Filter

System Integration

These filters were integrated into a micropower analog bionic ear[9].
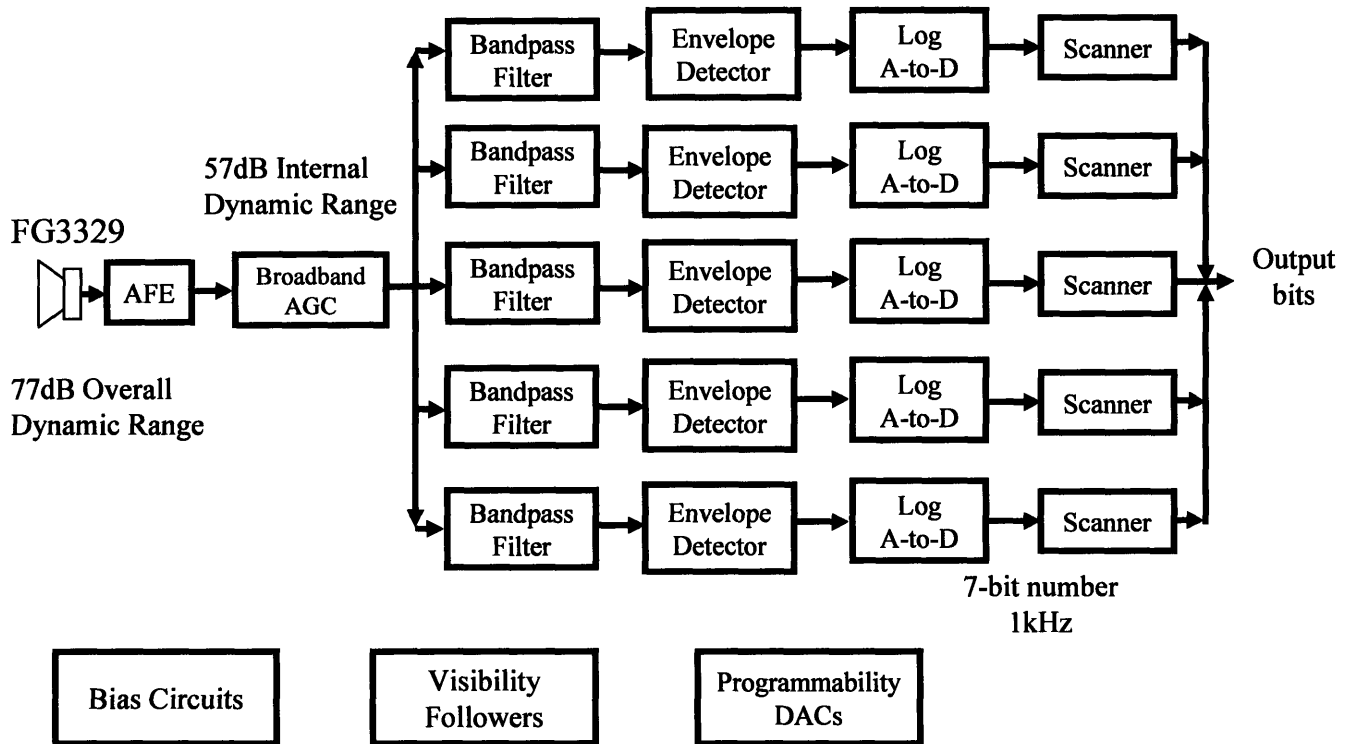
Figure 2.21: Bionic Ear Topology

Conclusions

Moving bionic ears from systems that are partially worn outside the body to true implants that are fully embedded inside the body requires overcoming a number of technical obstacles. One important obstacle involves ways to substantially lower the power of the signal processing. A first step in that direction is exploring and customizing filter technologies for these unique requirements. These filters must operate on low voltage rails, with large dynamic range, be tunable over six octaves of frequency, and use minimal power. This work has demonstrated that subthreshold Gm-C filters with explicit capacitive attenuation can meet those specifications while keeping harmonic distortion at low levels over the entire operating regime.

References

[1]F. A. Spelman, "The Past, Present, and Future of Cochlear Prostheses," *IEEE Engineering in Medicine and Biology.* May/June 1999 p.27

[2]Philipos C. Loizou, "Mimicking the Human Ear," *IEEE Signal Processing Mag.,* vol. 5, pp. 101-130, Sept. 1998.

[3]R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A low-power wide-dynamic-range analog VLSI cochlea," *Analog Integrated Circuits and Signal Processing,* Vol. 16, pp. 245-274, 1998.

[4]R. J. W. Wang, R. Sarpeshkar, M. Jabri, and C. Mead, "A Low-Power Analog Front-end Module for Cochlear Implants," presented at the XVI World Congress on Otorhinolaryngology, Sydney March 1997

[5]W. Germanovix and C. Tourmazou, "Design of a Micropower Current-Mode Log-Domain Analog Cochlear Implant," *IEEE Trans.on Analog and Digital Signal Processing,* vol. 47, pp.1023-1046, Oct. 2000.

[6]R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A low-power wide-linear-range transconductance amplifier," *Analog Integrated Circuits and Signal Processing,* Vol. 13, No. 1/2, May/June 1997.

[7]E. Sanchez-Sinencio and J. Silva-Martinez, "CMOS Transconductance amplifiers, architectures and active filters: a tutorial," *IEE Proc.-Circuits Devices Syst,* vol. 147, pp.3-12, Feb. 2000

[8]W. M. Siebert, *Circuits, Signals, and Systems.* Cambridge, MA: The MIT Press, 1986.

Chapter 3 : Jump Resonance[*]

Jump resonance is the phenomenon in nonlinear filters where the amplitude of a sinusoidal output changes discontinuously near the peak of its frequency response for a smooth change in the input amplitude or frequency. It happens in high-Q filters when the signal amplitude is large. An example of this phenomenon is shown in Fig. 3.1 where the output amplitude of a filter is plotted as the input frequency is swept up and then down. As the frequency is swept upward, the amplitude slowly increases and then jumps. When the frequency is swept back downward, the amplitude continues to vary smoothly past the frequency where the amplitude jumped up and jumps down at a lower frequency. These jumps and the hysteresis between them are the phenomenon referred to as jump resonance.

In the past fifty years, jump resonance has been studied in a variety of systems [1]. Control systems develop jump resonance when the actuator limits [2, 3]. Filters develop jump resonance when the signal passes beyond the linear range of the filter [4-6]. The theory of much of this work was captured by Fukuma and Matsubara when they developed a criterion based on describing-function analysis for predicting the onset of jump resonance [7]. The standard technique has been to design a static system that only demonstrates jump resonance for signal levels beyond its operating range.
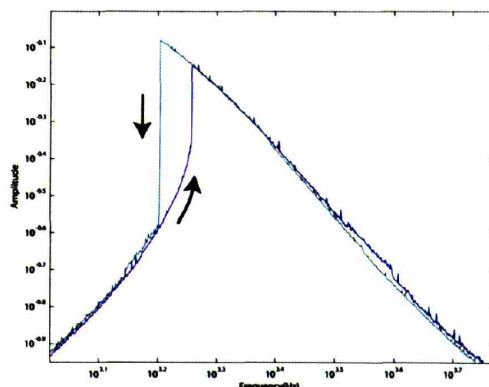
Figure 3.1: Experimental Example of Jump Resonance.

This chapter explores other methods for dealing with jump resonance. My application of interest is in developing micropower active analog filters for audio applications, particularly bionic ears (cochlear implants) where low-power operation and mimicking biological signal processing is important [8-11]. Low-Q active analog filters work well and rarely exhibit jump resonance. At high Q's such filters exhibit increased harmonic distortion, compression, a reduced instantaneous dynamic range, and a qualitatively new behavior, jump resonance, which is not a graceful degradation, but, a discontinuous one. Discontinuous behavior is not acceptable for several applications where active analog filters are used such as in a speech spectrum analyzer [12], a capacitive flow sensor [13], a lightning sensor [14], in very low frequency resonant sensor systems [15], and in bionic ear processors [9-11].
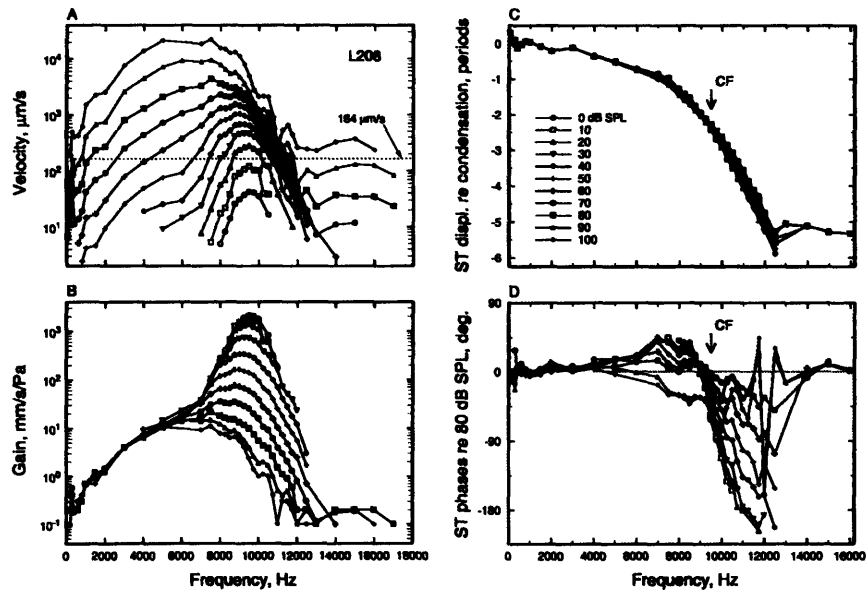
Figure 3.2: Basilar Membrane Measurements. Figure used with permission from Ruggero et. al. [17], copyright 2000, The National Academy of Sciences of the United States of America.

The biological cochlea is built with active analog filters with outer hair cells serving to perform amplification. It exhibits distortion and compression but jump resonance does not appear to have been reported in the cochlea [16]. Figure 3.2 shows that the motion of the basilar membrane within the cochlea encodes sound information and demonstrates an amplitude-dependent frequency response [17]; the data were taken in a rodent, the chinchilla. Figure 3.2B clearly demonstrates that both the gain and Q reduce in the cochlea as the signal amplitude

increases. I will show that, if the Q of the filter is appropriately lowered with amplitude, jump resonance can be avoided.

Other silicon systems have adapted their Q with signal amplitude: Some of these include circuit models of the cochlea, built from a cascade of filters. In these systems, the Q of each stage is adjusted to control the gain of the system and to recreate biological data [18, 19, 20]. A second set of systems uses bandpass filters to process audio, but adjust their Q to capture biological behavior [21-22]. These reports do not discuss the phenomenon of jump resonance because they did not try and operate at as high a Q as possible or were digital implementations.

It is worth noting that *any* active analog filter will benefit from adaptation in its Q with signal amplitude if it is to operate over a wide input dynamic range irrespective of whether its Q is increased to a high enough value to trigger jump resonance: At high Q, the minimum detectable signal of the filter is low but distortion occurs at relatively low input amplitudes while at low Q the minimum detectable signal is high but distortion occurs at relatively high input amplitudes. Thus, to operate over a wide input dynamic range, the Q must be high at small-signal levels and lowered at large-signal levels; otherwise, enormously more power is needed to operate the high-Q filter with a wide dynamic range. Hence, both in biology and in electronics, where power-efficient operation is important, active analog filters adapt their Q [20]. Like in an AGC [10, 11], the overall dynamic range of an adaptive-Q system is then wider than its instantaneous dynamic range. I add to the existent knowledge of adaptive-Q systems by showing how the adaptation in Q with signal amplitude can be architected to avoid jump resonance.

This chapter starts by presenting a simple Gm-C filter building block in Section II. In Section III, the Fukuma-Matsubara method is used on this filter to demonstrate the conditions for jump resonance. I develop intuition by considering different bias conditions with the method. In Section IV, I present my graphical method. This method accurately predicts not only the presence of jump resonance, but the actual value of the output amplitude for each input condition. I relate my method to feedback-system intuition on jump resonance. Then I show how the effects of changes in the design of the linear and nonlinear portions may be examined since they are represented in standard forms. In Section V, I demonstrate experimental

79

agreement of results from our chip with our graphical method and feedback intuition. I present a circuit for Automatic Q Control (AQC) in Section VI. It extends the range of jump free operation by creating a system that degrades in a manner analogous to the biological cochlea. I demonstrate my theory on an integrated-circuit realization of a reconfigurable filter system that was tested in open-loop configuration to verify our graphical method and in AQC mode to demonstrate the engineering solution. In Section VII, I conclude by using a state-space approach to analyzing the filter. This approach leads to useful connections to the Duffing equation and predicts both jump resonance and other nonlinear behavior not modeled by the describing function based methods discussed above.

Section 1. Filter Topology

Analog filters have been built in integrated circuits using a variety of technologies including: Active-RC, MOSFET-C, Gm-C, Log Domain, and Switched-Capacitor. Of these, Gm-C designs are among the most power efficient because they directly use the buffering and transconductance of the transistors. The Active-RC, MOSFET-C, and Switched-Capacitor designs all require an amplifier with significantly more bandwidth than the filter itself and high gain to ensure linear operation leading to relatively less power-efficient solutions. The trade-off of using the transconductance directly is that nonlinear operation becomes more significant. This combination of power efficiency and nonlinearity makes the filter in Fig. 3.3, the two-integrator biquad [23], a good example for our analysis.
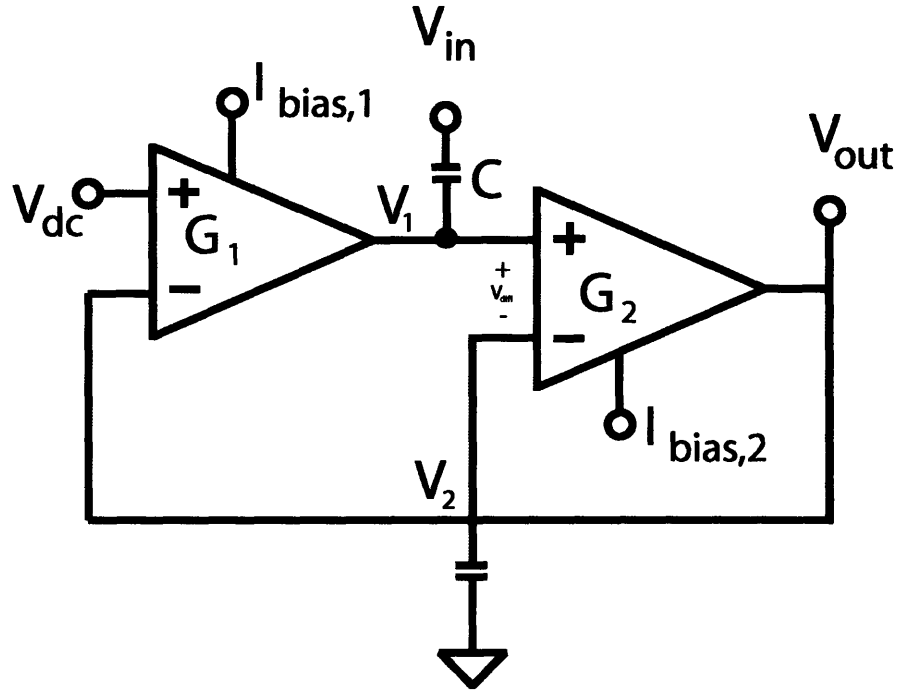
Figure 3.3: Two-Integrator Biquad.
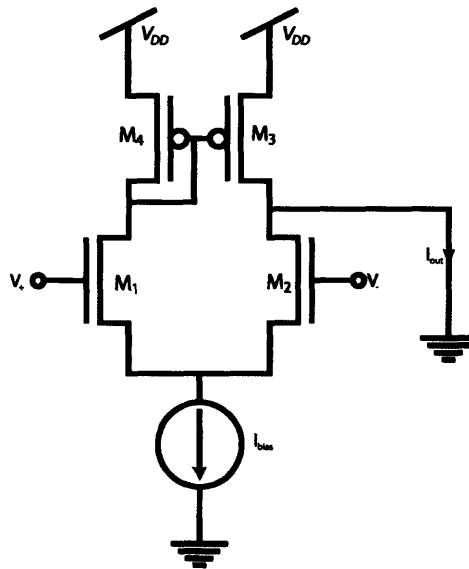
Figure 3.4: Transconductor Implementation.

Our GM-C filter is built from two basic building blocks. The capacitors integrate the current applied to them to create a voltage. Poly-silicon or metal-insulator-metal capacitors in integrated circuits can be very linear across the voltage range of the process and are assumed linear throughout this paper. The current that they integrate comes from differential

transconductors which can be nonlinear. For linear analysis, the output current is defined as the difference between the input voltages scaled by a transconductance factor. The nonlinear analysis requires a large signal model of the transconductor. There are many ways to build these transconductors with different topologies offering different large-signal behavior, but a common basic design is the one shown in Fig. 3.4. When it is biased for subthreshold operation, its large signal behavior is given by [24]:

$$I_{out} = I_{bias} \tanh\left(\frac{\kappa}{\phi_t}\left(\frac{V_+ - V_-}{2}\right)\right) \tag{3.1}$$

where κ is the process dependent subthreshold exponential constant and ϕ_t is the thermal voltage.

The analysis of the filter of Fig. 3.3 begins with finding its linear transfer function using linear transconductor models. For small signals with a saturating nonlinearity, like the tanh above, the filter behaves linearly with a transfer function given by

$$H(s) = \frac{V_{out}}{V_{in}} = \frac{s g_{m2} C}{g_{m1} g_{m2} + s g_{m2} C + s^2 C^2}, \tag{3.2}$$

where g_{m1} is the transconductance of G_1 and g_{m2} is the transconductance of G_2. It is a bandpass filter with programmable center frequency and quality factor. Since the nonlinear behavior is due to the transconductors, the signal amplitude at their differential inputs is important. The relative amplitude of these transconductor inputs when the system first becomes nonlinear can also be found by linear analysis. The transfer function to the differential input of the first transconductor has the same magnitude as the transfer function to the output. The maximum magnitude of this transfer function is 1, when the two real components in the denominator of Eq. 3.2 cancel each other out (at $s = j\sqrt{g_{m1} g_{m2}}/C$.) The transfer function to the differential input of the second transconductor, V_{diff}, is a resonant highpass transfer function referred to as $H_{diff}(s)$ in this paper. By writing the equation in terms of τ and Q

82

below, it is clear that the maximum magnitude of this node is very near Q and the radial frequency at the maximum, ω_{max}, is very near $1/\tau$.

$$\frac{V_{diff}}{V_{in}} = \frac{s^2 C^2}{g_{m1}g_{m2} + sg_{m2}C + s^2 C^2} = \frac{s^2 \tau^2}{s^2 \tau^2 + s\dfrac{\tau}{Q} + 1} = H_{diff}(s) \tag{3.3}$$

$$\tau = C \Big/ \sqrt{g_{m1}g_{m2}} \tag{3.4}$$

$$Q = \sqrt{g_{m1}/g_{m2}} \tag{3.5}$$

For high-Q systems, where jump resonance is important, the second transconductor has a much larger differential input signal than the first transconductor and only its nonlinearity needs to be analyzed. In this paper, we assume that only the second transconductor, G_2 of Fig. 3.3, is nonlinear. The experimental agreement shown later demonstrates that this assumption holds even for very large signal swings.

Nonlinear behavior is added to the model by using describing functions, a method proposed by Goldfarb in the 1940's [25]. The essence of this analysis is that many nonlinear systems can be analyzed by only considering signals at the frequency of the input. The nonlinearity changes the amplitude at this frequency and creates additional frequency components. If the linear portion filters out these other frequencies, the nonlinear portion can be modeled as a linear system that changes its parameters with input amplitude. More recently, more general methods have been proposed that model the limited ability of the linear blocks to filter out higher harmonics [26]. Inclusion of those techniques was not found to be necessary for this work but would definitely be of interest for more accurate extensions of it.
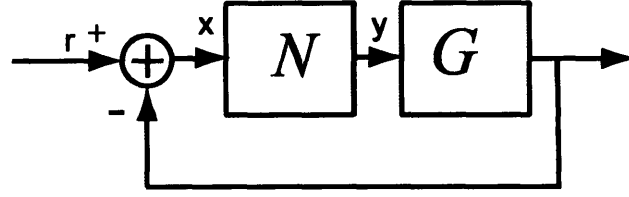
Figure 3.5: Describing Function Block Diagram.

To fit the formalism that Goldfarb used when he proposed the method, the system is rewritten in the form of the block diagram of Fig. 3.5 [25]. This requires reorganizing the block diagram representing the circuit. A traditional block diagram is built to show the signal input, V_{in}, and the signal output of the filter, but for nonlinear analysis those signals are secondary to the division of the system into linear and nonlinear blocks. The inputs and outputs of the filter, as it will be used in a bigger system, can be calculated within the linear block. So the system is represented by just two blocks: one linear and one nonlinear. Since the second transconductor is the only assumed nonlinearity, we combine the linear elements of the loop around the nonlinearity to form the loop transmission G starting at V_{diff} :

$$G = \frac{g_{m2}\left(g_{m1} + sC\right)}{s^2 C^2} \tag{3.6}$$

The non-linear block, N, is the tanh from Eq. 3.1.

Describing function analysis only considers the fundamental frequency, so it is important that the higher harmonics that are introduced by the nonlinearity are filtered by the linear portion. In our case, the lowpass frequency response of G ensures this behavior.

Section 2. Fukuma-Matsubara

Fukuma and Matsubara developed a criterion for jump resonance by performing algebraic manipulations on an equation for the derivative of the amplitude response [7]. The steps of their derivation are included as Appendix A of this paper for completeness. For comparison to our method, we point out a few key features of their technique. In order to separate the linear and nonlinear terms of the system, they define a new function H derived from the loop transmission G in Eq. 3.6. It is then decomposed into real and imaginary parts.

84

$$H = \left(G(j\omega)\right)^{-1} = h_r(j\omega) + jh_i(j\omega) \tag{3.7}$$

They demonstrate that a necessary and sufficient condition for jump resonance is the presence of this function H in a region of the complex plane defined by the nonlinearity and the signal amplitude. For a piecewise linear, saturating nonlinearity, the forbidden region is defined by a series of circles, a selection of which are shown in Fig. 3.6. Each circle represents the forbidden region for a different signal amplitude.
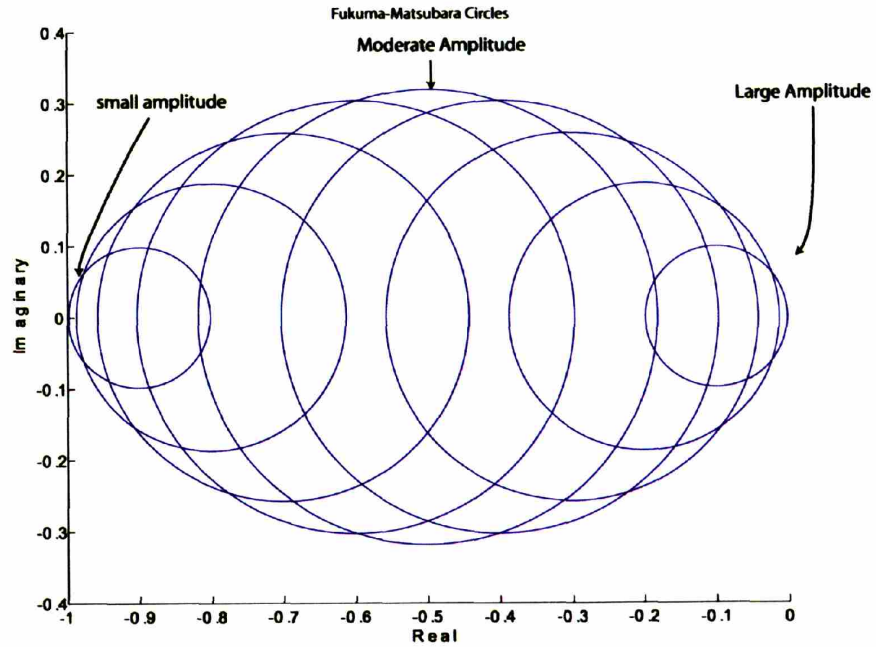
Figure 3.6: Fukuma-Matsubara Circles.
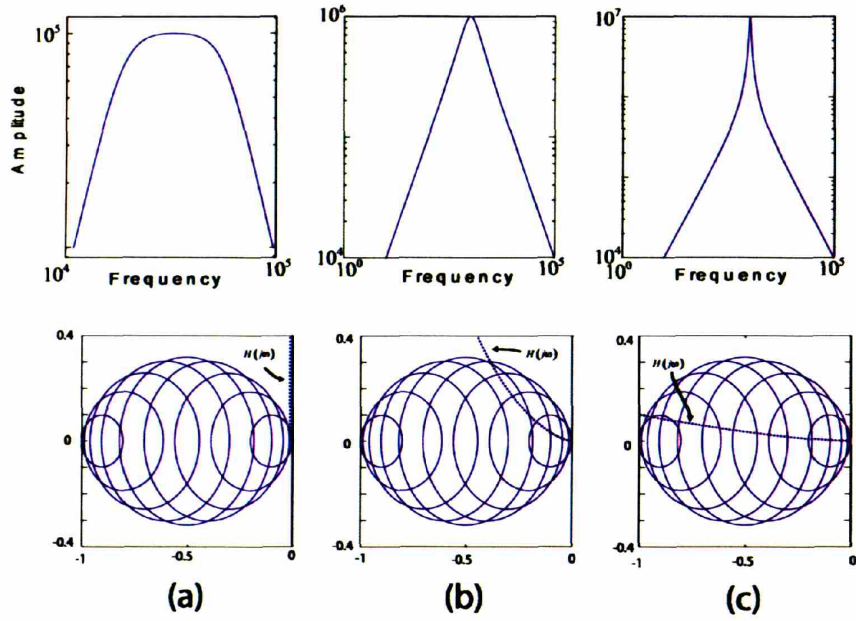
Figure 3.7: Fukuma-Matsubara Plots with Different Values of Q.

The analysis in Section II showed that as Q increases, the system becomes nonlinear for smaller input amplitudes and therefore is more likely to exhibit jump resonance. This intuition can be verified by plotting the frequency response of H for three different values of Q in the Fukuma-Matsubara plots of Fig. 3.7. The low Q system of Fig. 3.7(a) only passes through a circle corresponding to very large amplitudes, while the high Q system of Fig. 3.7(c) passes through all of the circles indicating jump resonance for all of the plotted amplitudes. Figure 3.7(b) shows a case where jump resonance is indicated for large amplitudes where H passes through the corresponding circles, but not for smaller amplitudes where it does not.

Section 3. Graphical Method

The Fukuma-Matsubara method predicts if jump resonance occurs, but it leaves many other questions unanswered. How does the peak frequency change? How big is the jump? And what is the shape of the new frequency response? Equally important, because it defines behavior in the inverse complex plane, it is difficult to develop intuition. Solving for the describing function solution addresses these issues [4].

There are two ways to find the describing function solution. Hiser used an iterative algorithm to solve for the operating point at each frequency [4]. His method starts with a low frequency input where the system has the same input-output relationship as the linear system. The linear model is used to calculate the signal amplitude at all the nodes of the filter. These signal amplitudes are then used to calculate a Distorted Transfer Function (DTF) using the nonlinearity. The DTF is used to calculate the amplitude at the next frequency and the steps are repeated.

We use a graphical approach that analyzes each frequency point independently. The system is divided into a parameterized linear system, $H_{diff}(s, g_{m2})$, and a nonlinear system, $g_{m2}(V_{diff})$. The parameterized linear system generates all of the same node amplitudes as the physical system for a given input condition. It is a parameterized linear system because a parameter corresponding to transconductance g_{m2} varies with the amplitude of the signal at the differential input, V_{diff}, of the nonlinear transconductor G_2. The nonlinear system sets a g_{m2} based on its input amplitude, V_{diff}, and the linear system sets a V_{diff} based on this g_{m2}. If the V_{diff} versus g_{m2} curves for the linear and nonlinear systems are plotted, the possible operating points for the physical system are points where the two curves intersect and yield a consistent solution.
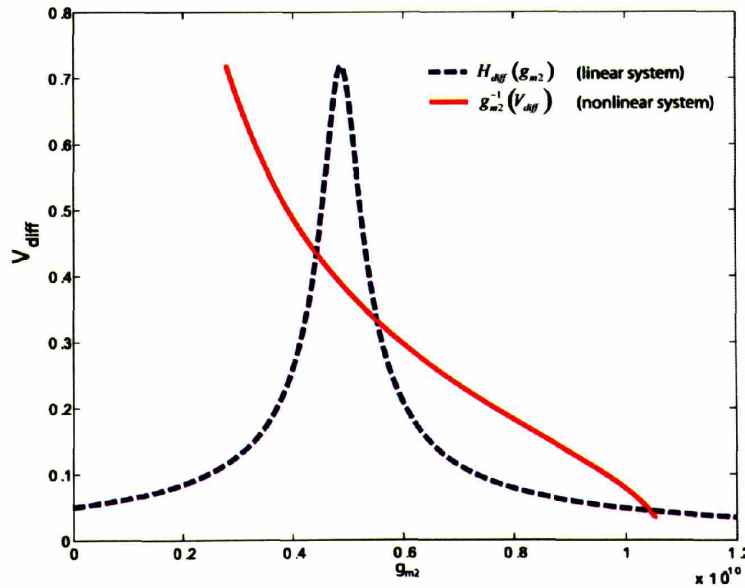
Figure 3.8: Graphical Operating Point Determination.

87

Figure 3.8 shows an example of one of these plots. The dashed curve is defined by the linear portion. It is given by:

$$V_{diff} = \left| V_{in} \bullet H_{diff}(s) \right|_{g_{m2}=g_m} = \left| V_{in} \bullet \frac{s^2 C^2}{g_{m1}g_m + sg_m C + s^2 C^2} \right|_{g_m}$$

$$= V_{in} \frac{\sqrt{\left(\omega^2 C^2\right)^2}}{\sqrt{\left(\omega^2 C^2 - g_{m1}g_m\right)^2 + \left(\omega C g_m\right)^2}} \tag{3.8}$$

This has a similar form to the magnitude response of the system and some of the intuition from that domain can transfer, but g_m is the varying quantity and ω is fixed. The numerator can be thought of as a constant in "g_m space" because the $\omega^2 C^2$ real term becomes a constant at a set frequency. The denominator cannot be placed in an equivalent pole format because the real component has a linear dependence on g_m.

The solid curve is a plot of the describing function of the nonlinearity, the tanh from Eq. 3.1.

$$g_{m2} = \frac{\int_0^{2\pi} \sin\theta \tanh\left(V_{diff} \sin\theta\right) d\theta}{2\pi V_{diff}} \tag{3.9}$$

It is obtained by taking the ratio of the amplitude of the fundamental output component to the amplitude of the sinusoidal input. In the plot of Fig. 3.8, we plot $g_{m2}^{-1}\left(V_{diff}\right)$ rather than $g_{m2}\left(V_{diff}\right)$, as in a standard describing function plot, so that the axes of the linear and nonlinear plots are consistent to enable the determination of their intersection. The values of a parameter α are included in the legend for discussion later in this section.

88

An example with an input that creates multiple stable operating points for the system is shown in Fig. 3.8. The intersection point on the right represents a high transconductance with small amplitude. The point on the left represents a low transconductance with large amplitude. There is also an intersection point with intermediate amplitude. In this example, the intermediate operating point is unstable and the other two operating points are stable.
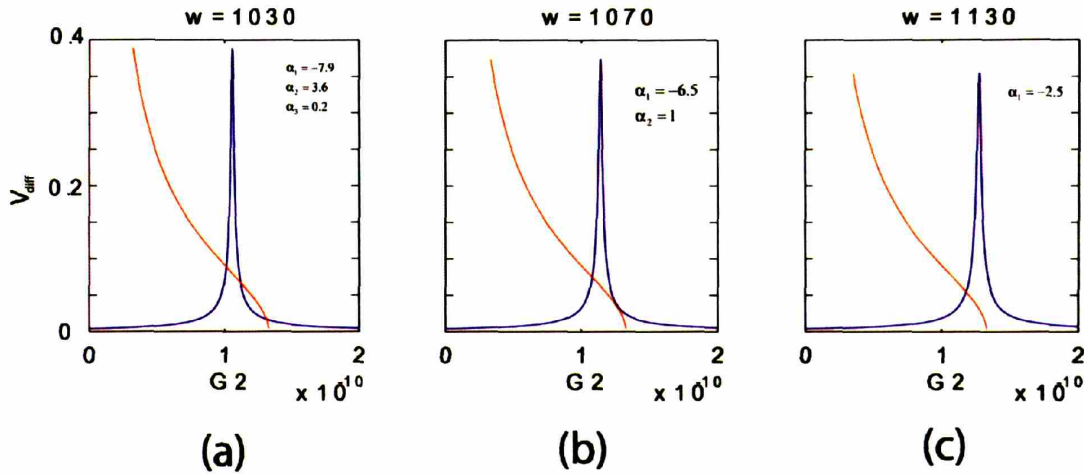
Figure 3.9: Jump Resonance in the Graphical Method.

This paper started with a demonstration of the frequency sweep behavior of this system with jump resonance. To see that behavior with the graphical method, multiple graphs are created, one at each frequency point. Although this seems tedious, intuitively one may just imagine the motion of the graphs to predict the system's behavior. Three of these graphs are shown in Fig. 3.9. The system starts at a low frequency for Fig. 3.9(a) with the same conditions as in Fig. 3.8. If the system is swept from lower frequencies, it will be at the high transconductance operating point because that is the point continuous with the linear operation at lower frequencies. The frequency is raised for Fig. 3.9(b). The nonlinear curve is in the same place as before, but the linear curve has changed. To a first approximation the linear curve simply shifts to the right. The high transconductance operating point has moved to slightly larger signal amplitude. Figure 3.9(c) shows what happens as the frequency increases further. The high transconductance point vanishes completely and the system jumps to the low transconductance point. This jump is the cause of the amplitude jump that is observed in jump resonance.
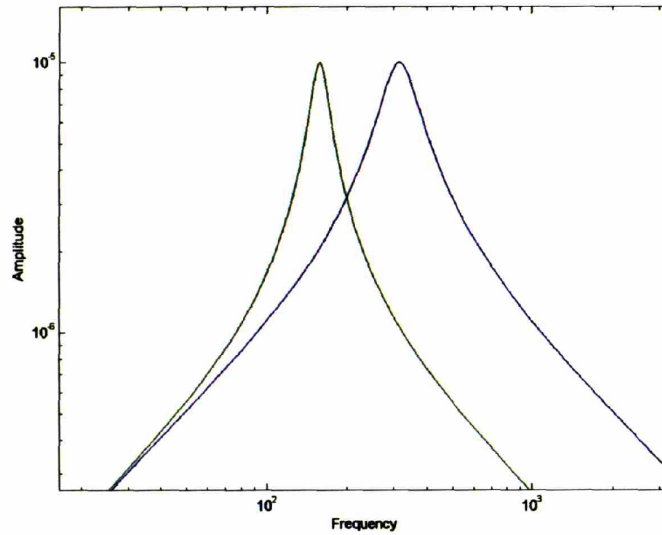
Figure 3.10: Frequency Response for Different Values of g_m.

The behavior can be equivalently visualized as Hiser proposed by plotting the frequency sweep of the parameterized linear system for different values of transconductance [4]. Two extreme examples of the frequency responses are shown in Fig. 10. Starting with a low frequency signal, the two curves are on top of each other. As the frequency increases, the signal amplitude at the input of the nonlinearity increases. This increase in amplitude decreases the transconductance, which lowers the center frequency of the filter, increasing the amplitude further. This positive-feedback loop is the step shown in the graphical method when the right stable operating point and the intermediate unstable operating point get closer to each other as the frequency is changed and eventually annihilate each other. After the center of this lower peak is passed, a negative-feedback loop is created: As the amplitude decreases, the center frequency moves back up, increases the amplitude of the signal, and opposes its original decrease.

90
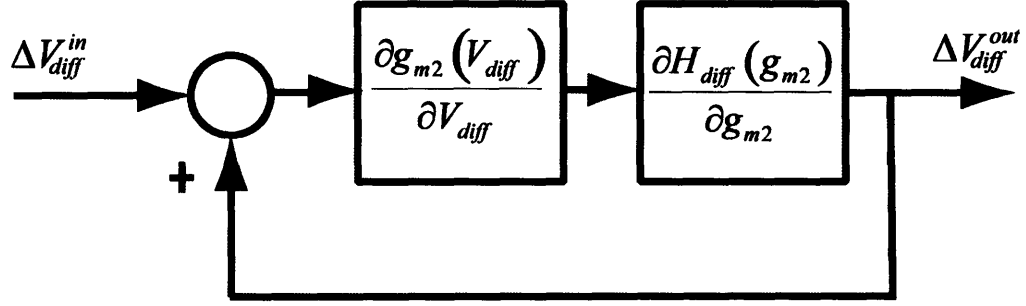
Figure 3.11: Loop gain Block Diagram.

These two methods can be united and quantified by calculating the loop gain of the loop created between the parameterized linear system and the nonlinear system represented by block diagram in Fig. 3.11. The small signal gains of the two labeled blocks are the slopes of the $g_{m2}\left(V_{diff}\right)$ curve and the slope of the $H_{diff}\left(g_{m2}\right)$ curve in Fig. 3.8. The loop transmission is then given by α in

$$\alpha = \frac{\partial g_{m2}\left(V_{diff}\right)}{\partial V_{diff}} \frac{\partial H_{diff}\left(g_{m2}\right)}{\partial g_{m2}} \tag{3.10}$$

That is to say, the slope of the linear curve is multiplied by the reciprocal of the slope of the nonlinear curve. At a point where the two curves are tangent, the slopes are equal and the positive-feedback loop gain is 1, causing the jump. This agrees with the intuition of the two curves moving away from each other destroying the stable operating point. The values of α are plotted in Fig. 3.12 for the two stable operating points as we sweep the frequency. Operating points, such as the middle operating point in Fig. 3.9(a) are unstable because they have a positive loop gain greater than 1 and are not plotted in Fig. 3.12.
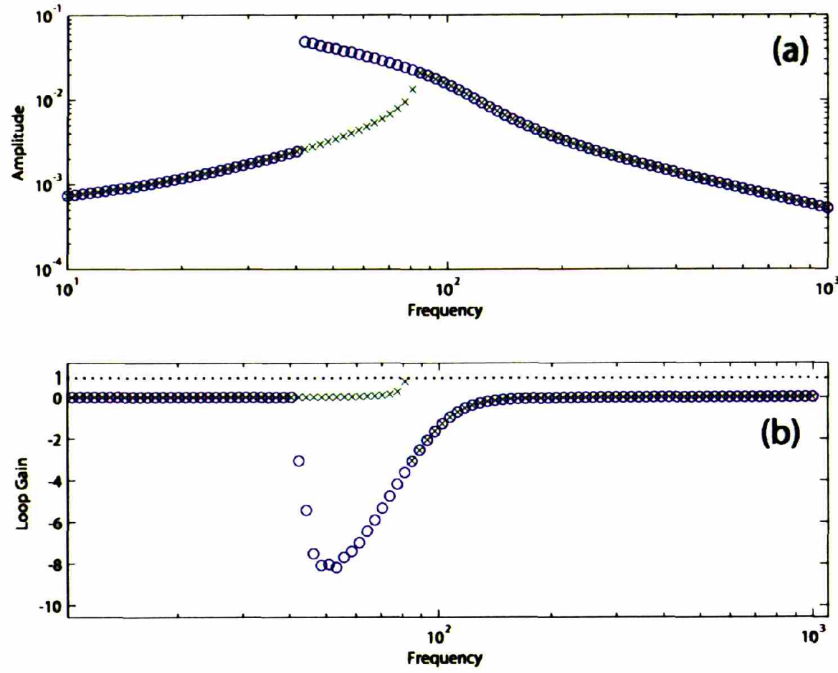
Figure 3.12: Loop Gain Plot

In Fig. 3.12, the x's represent the operating point with a smaller signal amplitude and the o's represent the operating point with larger amplitude. Both sweeps in the rising and falling direction have a loop gain of 1 at the jump. But, as Hiser suggested the sweep upward is characterized by positive feedback. The distinction between positive and negative feedback can be visualized by considering the slopes of the two lines at the points of intersection. When both curves have the same sign for the slope, as in the right most intersection point of Fig. 3.9(a), that is positive feedback. The downward sweep is largely negative feedback since the single stable operating point is defined by the intersection of curves with opposite slopes. As we lower the frequency again, we quickly sweep the operating point past the peak of the $H_{diff}\left(g_{m2}\right)$ curve and transition from negative to positive feedback, and the curves intersect at a new operating point higher up the $H_{diff}\left(g_{m2}\right)$ curve where α is again 1 at jump resonance, and less than 1 at the new family of operating points beyond the new jump resonance. Hysteresis is created because the two jump resonance points where $\alpha=1$ are not the same in the upward and downward direction.

92
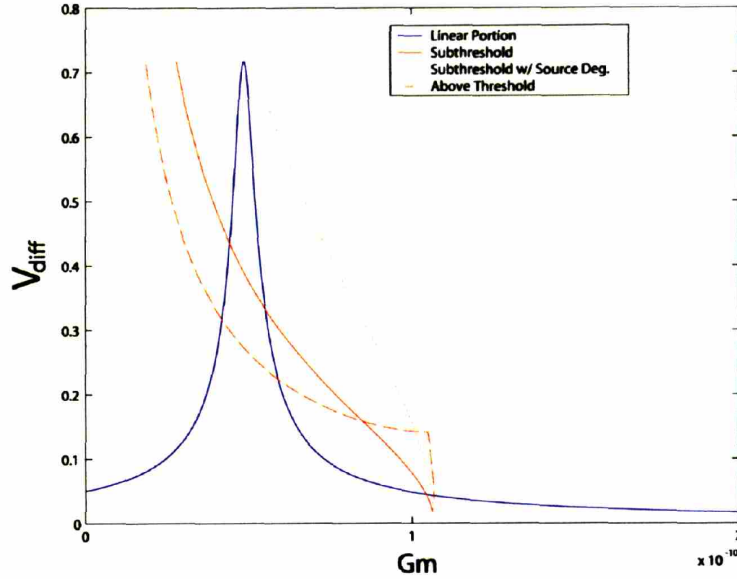
Figure 3.13: Jump Resonance from Different Transconductors.

The intuition from these two views of the system, the frequency point and swept frequency view, creates an opportunity to explore ways to change the jump resonance behavior of the system. The effect of the nonlinearity is now directly related to the describing function. In this example, the behavior is set primarily by the $1/V_{diff}$ behavior of the saturating nonlinearity for large signals. Figure 3.13 shows the effect of three different transconductors, with the linear range of the above-threshold transconductor adjusted to have a similar scale. The solid line is the simple OTA that is used in this design. If source degeneration is used, the nonlinearity becomes the dotted line. This doubles the amplitude at which jump resonance occurs, but does not change the behavior qualitatively. The dot-dash lines show the effect of an above-threshold transconductor, with a scaled linear range. It is more linear overall, but once it becomes nonlinear it has a similar shape to the curve already studied.

Collectively, our analysis shows that feedback interaction between filter parameter changes due to nonlinearities and the signal amplitude changes which cause them create jump resonance. Jump resonance occurs when this positive feedback loop gain is one. Expanding the maximum signal swing of the transconductor improves the amplitude at which jump resonance occurs, but after those limits are reached, the jump resonance phenomenon is similarly manifested.

93

Section 4. Experimental Agreement

Qualitative agreement between the graphical method and the experimental data has been demonstrated, but our technique is not merely qualitative. It also gives quantitatively correct measurements. The two-integrator filter was fabricated on AMI's 1.5μm process through MOSIS. We used Matlab to calculate the stable points by creating the linear and nonlinear curves of Fig. 3.8 for each sinusoidal input signal. The simulations were all done with reference to V_{diff} but this is not a physical node in the circuit. So, linear transformations were used to convert this voltage to the output voltage: That is,

$$V_{out} = V_{diff} \frac{g_{m2}}{sC} \tag{3.11}$$

First, sweeps were taken with a lock-in amplifier. The transconductors were kept at constant bias currents while sine waves were applied to V_{in}. Three different amplitudes were used with a hundred frequency points for each amplitude. Because each frequency point was taken as an independent measurement, only the operating point with lower amplitude was measured at each frequency. Figure 3.14 shows simulations from the graphical method as thin lines with points for the measured values for the frequency sweep. The values of the capacitance, C, and the describing function of the nonlinearity were taken as the designed values, while the transconductances were varied to fit the smallest amplitude curve. The fits for the higher curves did not require any additional free parameters.
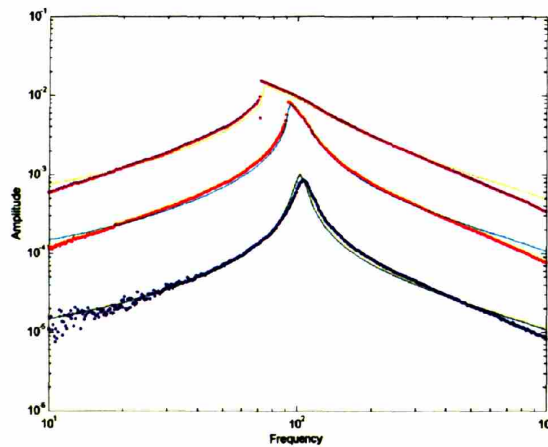
Figure 3.14: Experimentally Measured Frequency Sweeps and Simulation Results.

94
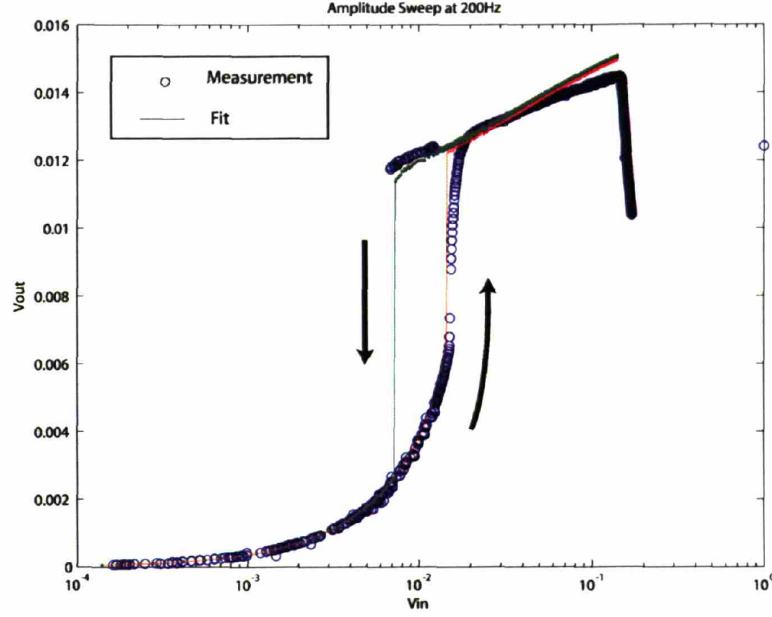
Figure 3.15: Experimentally Measured Amplitude Sweeps with Simulation.

Both stable operating points may be viewed by performing a continuous sweep. In Fig. 3.15, the frequency is kept constant and the input amplitude is varied. Again the fit is plotted as a line with the measurements as points. This measurement was taken as a true sweep, with a function generator continuously varying the amplitude of the sine wave and a spectrum analyzer measuring the signal amplitude such that the high and low stable points are measured in the hysteresis region. This plot is easier to calculate with the graphical method because the linear curve is simply scaled with the input amplitude to determine the intersection points with the nonlinear describing function curve. Again, good agreement is seen between the measurement and simulation.

Section 5. Automatic Q Control

One way to avoid jump resonance is to lower the Q of the filter for inputs with large amplitude such that the differential input, V_{diff}, to the transconductor, G_2, always stays in its linear regime. This can be done with the system shown in Fig. 3.16. The filter is the same as has been discussed so far. The envelope detector is a rectifier followed by a peak detector that was developed for low power wide-dynamic-range energy detectors and is described in [27]. The

95

output of the envelope detector is placed through a Q-control circuit which implements the relation:

$$Q_{out} = \frac{Q_{max}}{\frac{E^2}{A} + B} + Q_{min}$$

(3.12)

where A and B are constants. This control equation was chosen as the simplest relationship that can define a general Q compression. These four values allow the specification of minimum and maximum values of Q associated with arbitrary values of the envelope. The squared relation is chosen to vary the Q quickly enough to avoid jump-resonance. The Q control signal is then used to bias the filter. Figure 3.17 is a photograph of a chip die where the filter and other building blocks from Fig. 3.16 were fabricated.
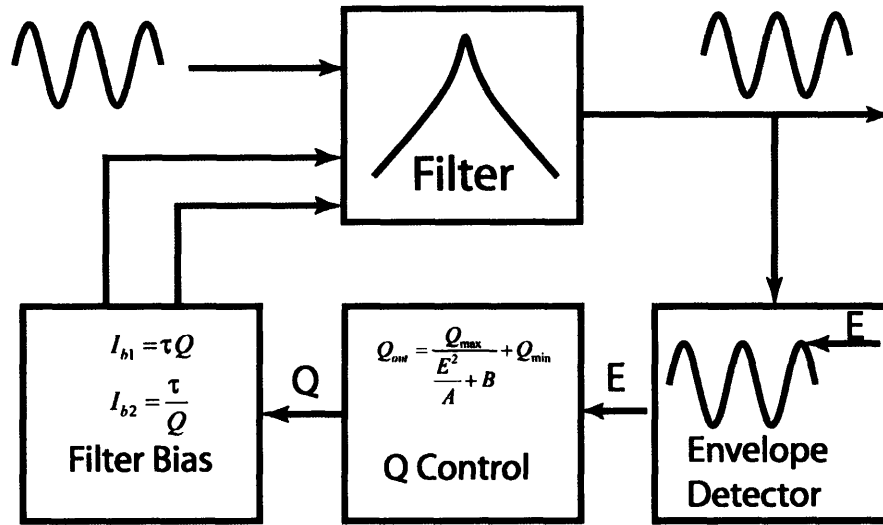
Figure 3.16: Feedback AQC System Block Diagram.
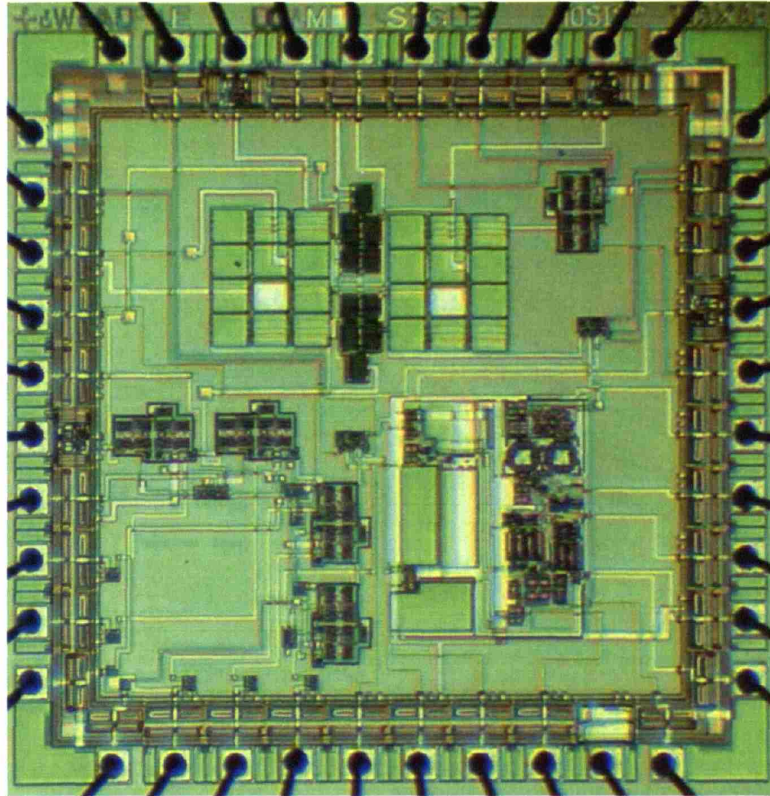
Figure 3.17: Die Photograph of AQC Feedforward and Feedback System.

Frequency sweeps on V_{in} at different amplitudes for the open-loop system are shown in Fig. 3.18. The system is linear for small amplitudes, but for larger amplitudes the center frequency shifts, the shape of the curve changes, and a jump is evident. The same sweeps for the feedback system in Fig. 3.19 show that the center frequency is held constant while the Q is lowered. This system has extended the range, but Fig. 3.20 demonstrates a problem that happens at larger amplitudes. The system is well behaved near the peak, but further away it shows steps. As the filter moves away from its peak, the envelope detector measures smaller amplitude and returns the Q to a high value, lowering the signal. Hence, the Q-adaptation creates the same sort of positive feedback that created the original jump resonance. In the example above, a single filter which is turned down to a very low Q, the discontinuities are evident; but, in a more complicated system they may not be because the effect is now away from the peak.
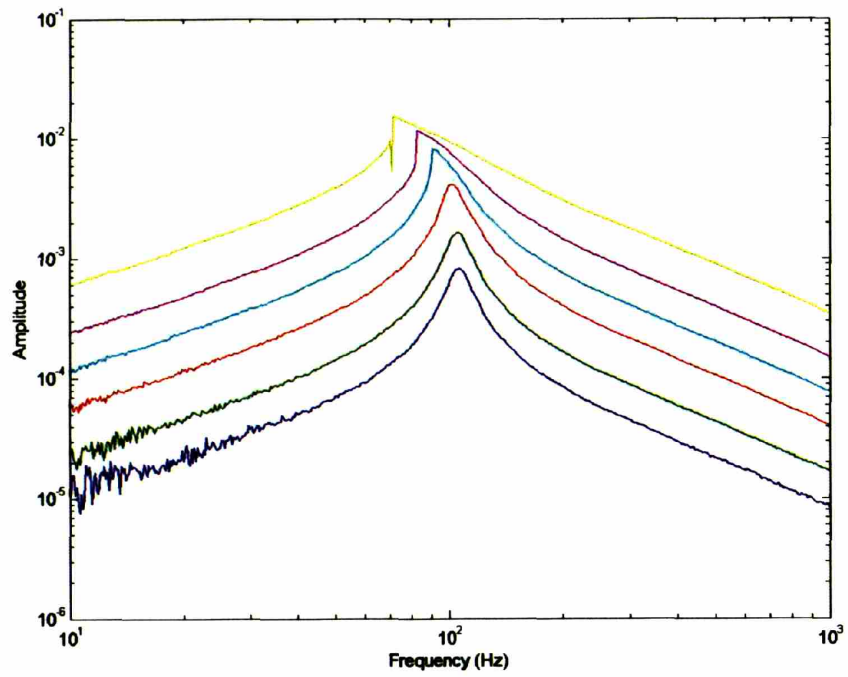
Figure 3.18: Measured Open Loop Frequency Sweeps.
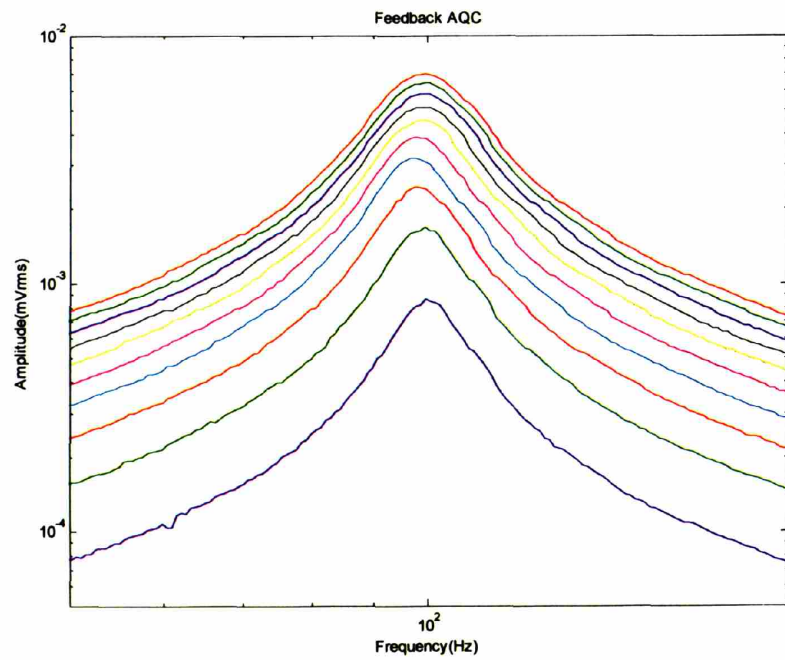
Figure 3.19: Measured Feedback AQC Frequency Sweeps.

98
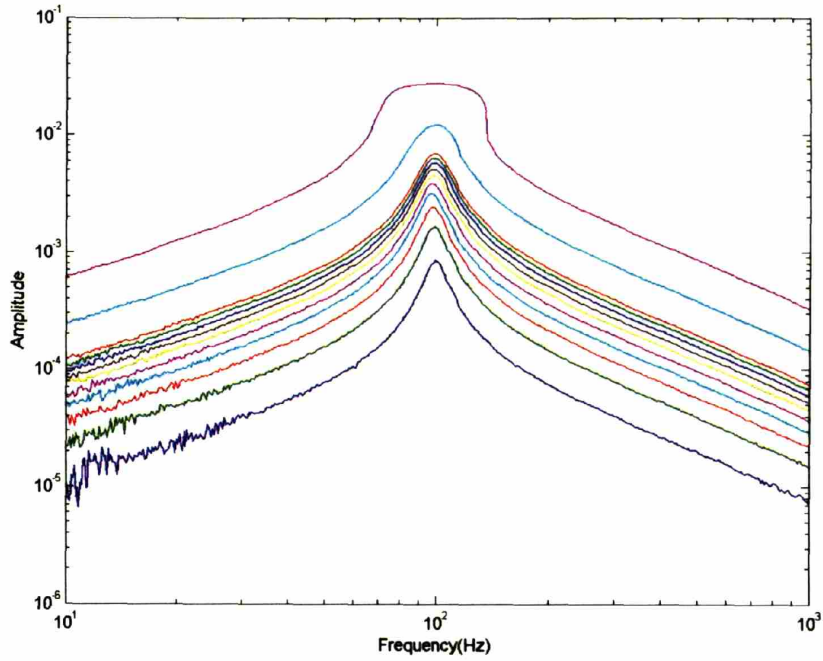
Figure 3.20: Measured Feedback AQC Frequency Sweeps with Extended Range.

These steps at non-resonant frequencies can be avoided entirely by going to the feed-forward system in Fig. 3.21. This system uses the same components as the feedback system and was tested as an alternative configuration of the chip shown in Fig. 3.17. With this configuration, the response of the system broadens smoothly as shown in Fig. 3.22. This Q-control strategy does not have any frequency selectivity, so it will lower the Q in response to any large input, even frequency components outside the frequency of interest. Frequency selectivity for the Q-control can be re-introduced by a low-Q pre-filter that has a wider dynamic range as a result of its lower Q. An example of such a system is discussed in [28].

99
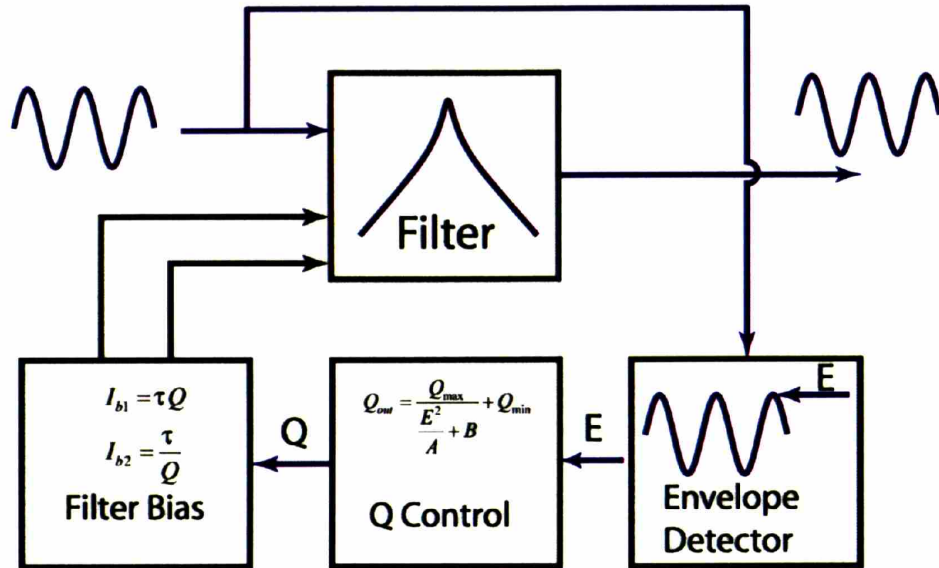
Figure 3.21: Feedforward AQC Block Diagram.
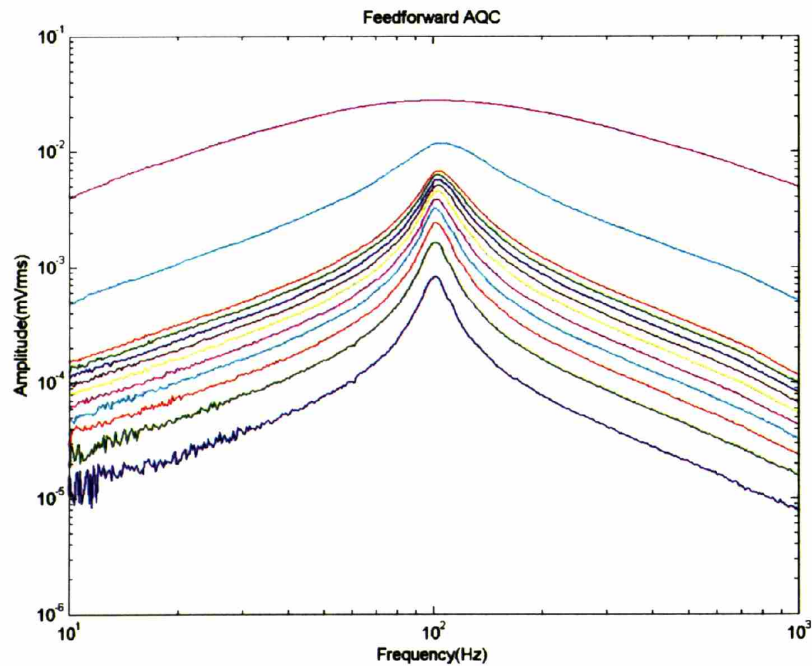
Figure 3.22: Measured Feedforward AQC Frequency Sweeps with Extended Range.

In looking at these logarithmic plots it can be difficult to judge the significance of this jump-free dynamic range extension. Comparing the power costs of alternative strategies provides a concrete measure. The dynamic range can be increased by either increasing the linear range or by decreasing the noise. It may be shown that, in thermal-noise limited designs, power and

100

dynamic range scale linearly whether transconductors with wide linear range are used or larger capacitors are used to lower the noise [20, 29, 30]. The dynamic range is defined as

$$DR = \frac{V_{max}^2}{2V_n^2} \qquad (3.13)$$

Comparing Figs. 3.18 and 3.20 we see that the maximum signal the system can handle is increased by a factor of 5 since the noise is relatively independent of Q [20]. From Eq. 3.13, then, our dynamic range extension is equivalent to reducing the power consumption by a factor of 25.

Section 6. State Space Analysis

The describing function methods presented in sections 2 through 5 accurately predicts the behavior of the system for many sinusoidal inputs. So, these methods are suitable for exploring jump resonance and the adaptive Q method of combating it. But, can jump resonance be avoided instead by pulling charge from one or more capacitors to force the system into a lower amplitude regime? To answer this question, an alternative method must be used. The filter can be modeled using differential equations for the capacitor voltages.

$$V_1 = \int \frac{-G_1 V_2}{C} + V_{in}(t) \qquad (3.14)$$

$$V_2 = \int \frac{G_2(V_1 - V_2)}{C} \qquad (3.15)$$

$$G_2 = \begin{cases} G_{2,0} & (V_1 - V_2) < V_{lin} \\ \dfrac{I_{max}}{(V_1 - V_2)} & (V_1 - V_2) \geq V_{lin} \end{cases} \qquad (3.16)$$

The state of the system is defined by V_1, V_2, and t. These equations were simulated in Matlab for a variety of initial conditions using parameters that lead to jump resonance. The trajectories led to the two attractors shown in Fig. 3.23 corresponding to the two stable amplitudes shown

101

earlier. The smaller amplitude trajectory has a high quality factor, Q, as demonstrated by its narrow shape. The larger amplitude trajectory has a lower Q.
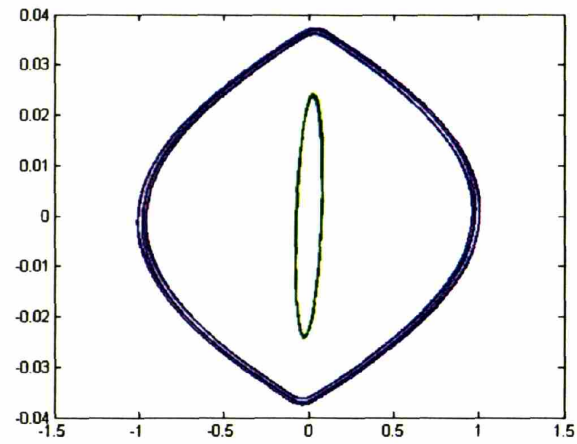
Figure 3.23: Jump Resonance Trajectories
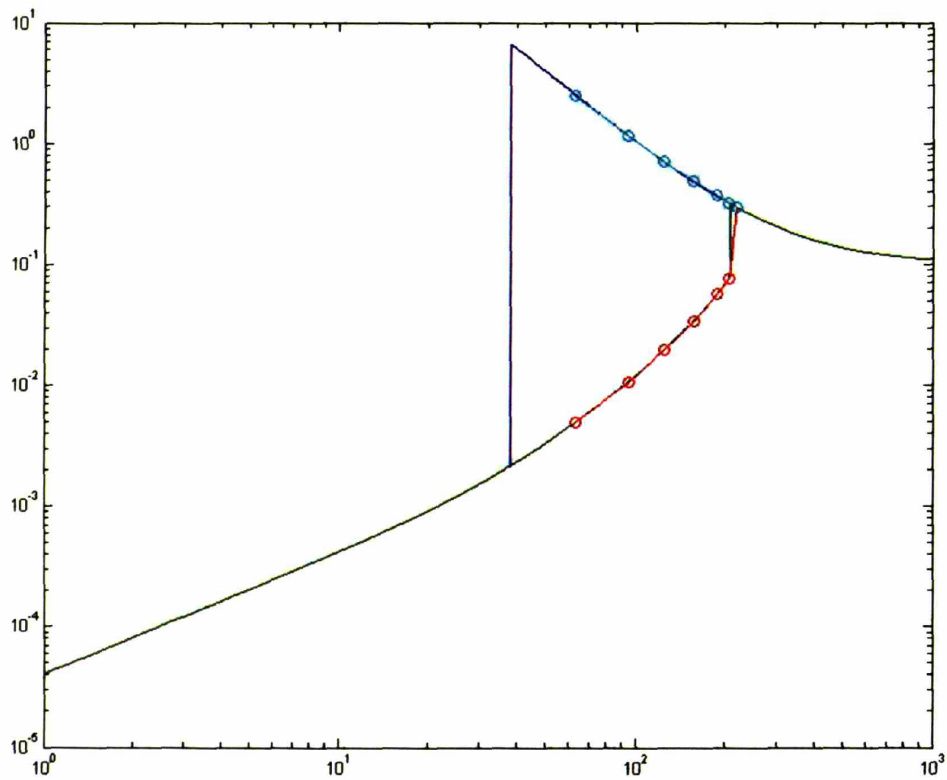
Figure 3.24: Jump Resonance Frequency Sweep with State Space Method.

This method was verified against the describing function method by plotting both the attractor amplitudes and the describing function behavior for a variety of frequencies in Fig. 3.24.

The possibility of eliminating jump resonance by changing the capacitor charge during operation, an instantaneous jump-resonance avoidance scheme, can be explored by determining the effect of initial conditions on final amplitude. In these simulations the system is driven with a sine wave

$$Vin(t) = A\sin(2\pi f t) \tag{3.17}$$

The system was given initial conditions for V_1 and V_2 when t=0 corresponding to a point on the V_1/V_2 plane. For each point, Eqs. 3.14 and 3.15 were simulated until a steady state solution was reached. The points were then color coded with the amplitude of the steady state solution. To increase computation speed, simulations were stopped for points that left the initial simulation region and those points were colored red. Figure 3.25 shows six of these domains of attraction plots for different frequencies. Figure 3.26 shows that the domains of attraction rotate around the DC value with phase of the driving sinusoid. In both figures, circles have been drawn at the point associated with the small amplitude trajectory and stars have been drawn at the point associated with the large amplitude trajectory.

(A) 5Hz   (B) 10Hz   (C) 15Hz

(D) 20Hz   (E) 25Hz   (F) 30Hz
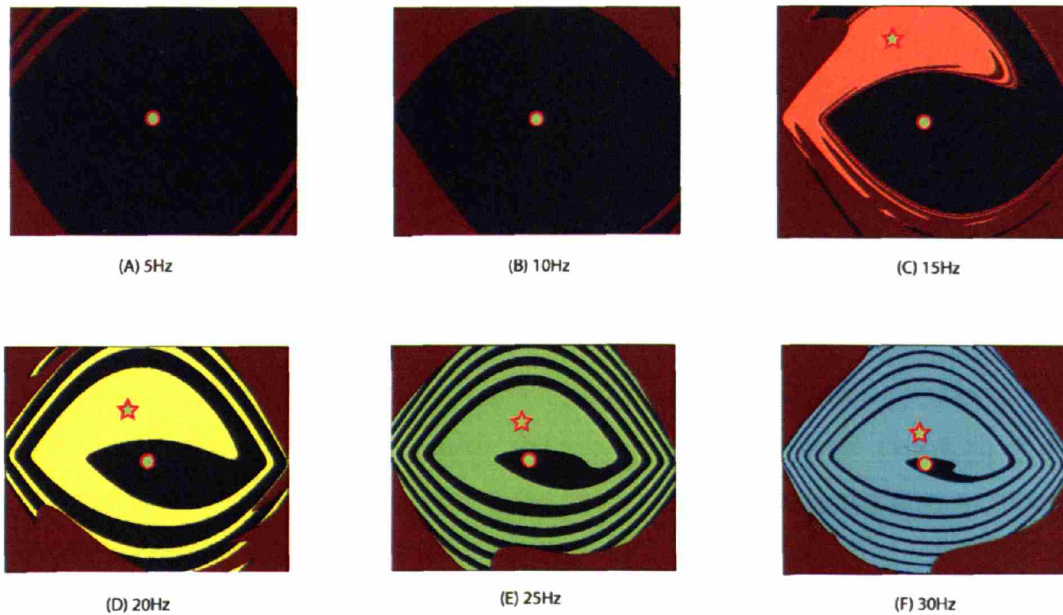
Figure 3.25: Domains of Attraction for Jump Resonance.

103
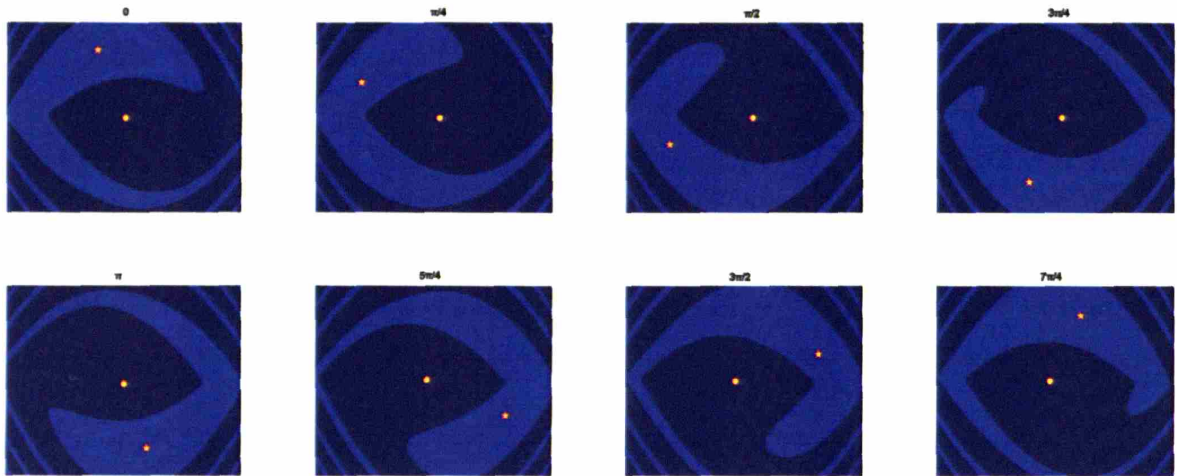
Figure 3.26: Domains of Attraction at 30Hz for Different Phase.

The plots of the domains of attraction complement the existing description of jump resonance. At small amplitudes, only the small amplitude attractor is visible. As the frequency increases, the large amplitude attractor appears. As the frequency increases further, the amplitude of the large amplitude attractor decreases and its range of attraction increases. The tails which circle the interior region may be of interest in the case of transients, but they are not reachable by steady state sinusoidal inputs.

Figure 3.26 can be used to analyze the instantaneous jump-resonance avoidance scheme. More than half of the charge stored on the second capacitor would have to be removed to force the system to move to the small amplitude trajectory. This would likely make the method of jump resonance avoidance impractical.
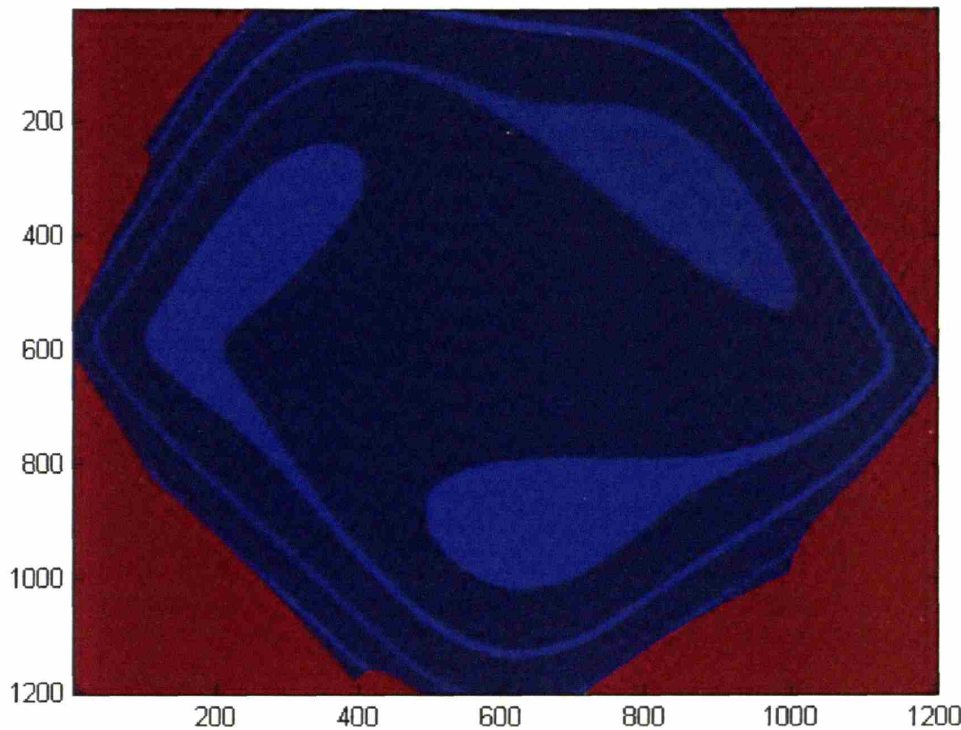
Figure: 3.27 Domains of Attraction at 50Hz.

The investigation of state-space methods can be continued, however. As the frequency is increased, another class of behavior emerges. Figure 3.27 demonstrates the domain of attraction plot at 50Hz. The three regions that surround the center of the plot are a subharmonic resonance. The output of the filter is periodic with a period thee times the input period. The same behavior is demonstrated in experimental measurements from a chip in Fig. 3.28.

The natural, or undriven, response of the system is a low frequency oscillation which slowly decays. This signal dominates in the bottom trace, V_2. In the middle trace, V_1, the low frequency tone is added to the higher frequency input signal. Because the transconductor connected to V_1 saturates for large signals, the system is only sensitive to the input during transitions of the low frequency response. The input provides the extra energy to sustain the low frequency oscillation for odd subharmonics.

105
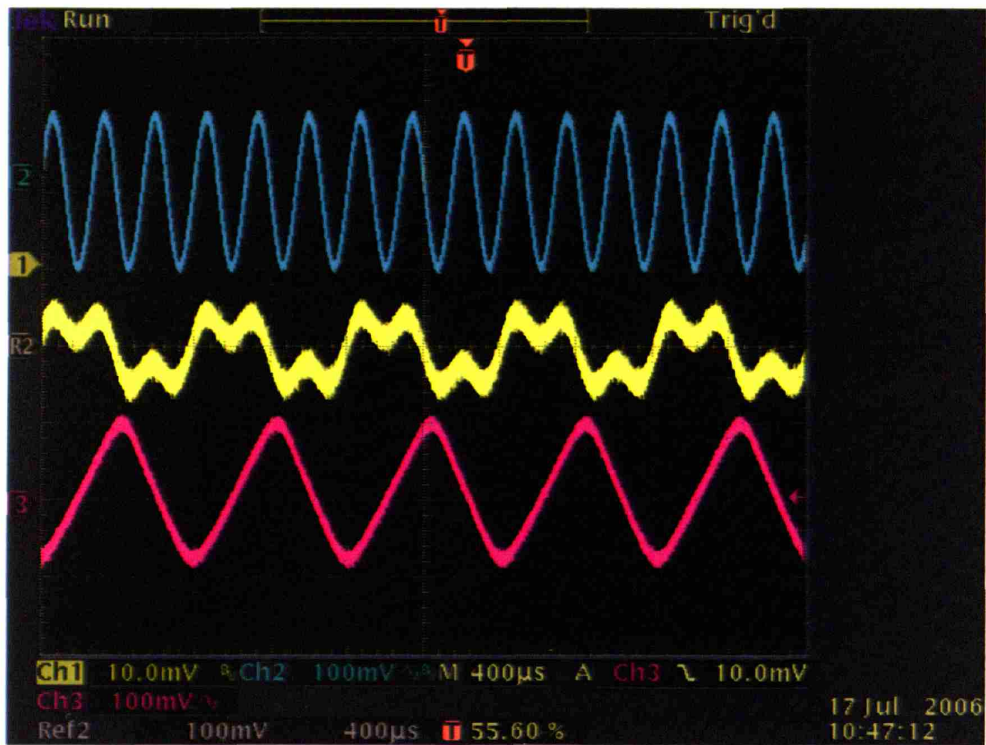
Figure 3.28: Experimental Measurement of Subharmonic Oscillations. The blue trace is the input signal. The yellow trace is V_1 showing first and third subharmonic content. The magenta trace,V_2, is predominately third subharmonic.
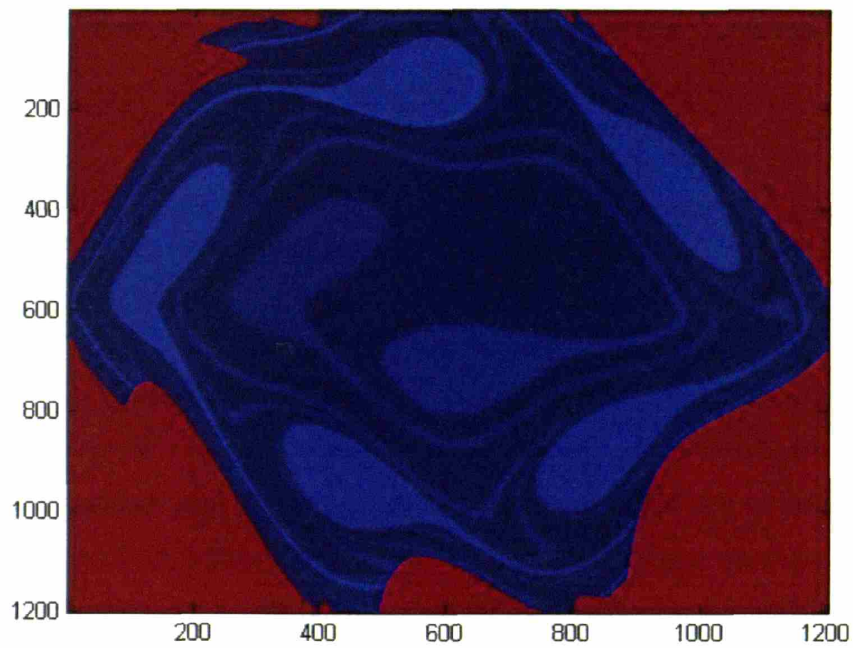
Figure 3.29 Domains of Attraction at 75Hz.

These subharmonics grow to include higher order modes at higher frequencies. At 75 Hz, both the 3rd order subharmonic and the 5th order subharmonic modes are present as shown in Fig. 3.29. These subharmonic trajectories are shown in Fig. 3.30.
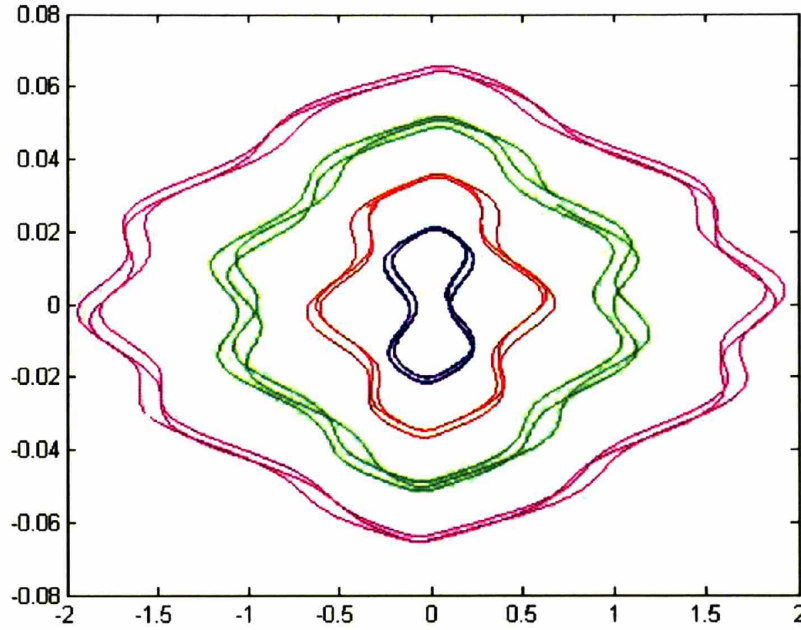
Figure 3.30 Trajectories of Subharmonic Modes.

Both this subharmonic behavior and jump resonance have been previously reported in a class of systems defined by the Duffing Equation [31]. Instead of approximating the nonlinearity directly as a simple function, this simplification approximates the effect of the nonlinearity as a simple function.

$$\frac{\partial^2 v}{\partial \tau^2} + k \frac{\partial v}{\partial \tau} + f(v) = B \cos v \tau \qquad (3.18)$$

Equations 3.14 through 3.16 can be approximated by Eq. 3.18 by first writing Eq. 3.14 and 3.15 as a single equation in terms of V_{in} and V_1.

107

$$V_2 = \frac{C}{G_1}\left(\dot{V}_{in} - \dot{V}_1\right) \tag{3.19}$$

$$\frac{C}{G_1}\left(\ddot{V}_{in} - \ddot{V}_1\right) = \frac{G_2}{C}\left(V_1 - \frac{C}{G_1}\left(\dot{V}_{in} - \dot{V}_1\right)\right) \tag{3.20}$$

$$\left(\ddot{V}_{in} - \ddot{V}_1\right) = \frac{G_1 G_2}{C^2} V_1 - \frac{G_2}{C}\left(\dot{V}_{in} - \dot{V}_1\right) \tag{3.21}$$

$$\ddot{V}_1 + \frac{G_2}{C}\dot{V} + \frac{G_1 G_2}{C^2} V = \ddot{V}_{in} + \frac{G_2}{C}\dot{V}_{in} \tag{3.22}$$

Equation 3.22 is a linear Duffing Equation for a constant value of G_2. The nonlinearity was introduced into the filter studied here by Eq. 3.16 which lowers the value of G_2 for large signals. This shows up in three terms in Eq. 3.22. The nonlinear behavior is primarily a concern in hiqh-Q systems, when the damping term is small, so the nonlinearity can be approximated as

$$\ddot{V}_1 + \frac{G_{2,0}}{C}\dot{V} + \frac{G_1 G_{2,0}}{C^2} V - \varepsilon \frac{G_1 G_{2,0}}{C^2} V^3 = \ddot{V}_{in} + \frac{G_{2,0}}{C}\dot{V}_{in} \tag{3.23}$$

Section 7. Conclusion

As input signals to a filter increase in amplitude, many nonlinear effects begin to occur: Harmonics are created, the fundamental amplitude fails to grow linearly with the input, and jump resonance develops. Jump resonance is of particular interest because it is a binary event that does not exist at small signals, but appears at larger amplitudes. This paper has presented a new graphical method for understanding and simulating this behavior that complements the standard Fukuma-Matsubara method. I have also shown that feedback interactions between filter parameter changes due to nonlinearities and the signal amplitude changes which cause them create jump resonance, i.e., jump resonance occurs when the positive feedback loop gain of this interaction is 1. The understanding developed from our analysis was used to develop a system that adapts its Q to avoid jump resonance. This system is not the solution for every filtering application, but when avoiding jump resonance is the limiting factor for input dynamic

range, our system can operate on 25 times less power than a traditional open-loop filter design. Finally, I have demonstrated an alternative analysis method based on state-space methods that can be used to evaluate non-sinusoidal signals and shown that the large body of work on the Duffing Equation is applicable to this problem.

Appendix A: Fukuma-Matsubara Method Derivation

The signals in Fig. 4 are defined:

$$r = R \sin \omega t \tag{A.1}$$

$$x = X \sin(\omega t + \phi) \tag{A.2}$$

The Nonlinear function is defined as:

$$N(X) = n_r(X) + jn_i(X) \tag{A.3}$$

Black's Formula gives:

$$Xe^{j\phi} = \frac{1}{1 + N(X)G(j\omega)} Re^{j\phi} \tag{A.4}$$

Solving for the magnitudes:

$$X|1 + N(X)G(j\omega)| = R \tag{A.5}$$

Substituting in for N:

$$X|1 + (n_r(X) + jn_i(X))G(j\omega)| = R \tag{A.6}$$

A new function H is defined:

$$H = (G(j\omega))^{-1} = h_r(j\omega) + jh_i(j\omega) \tag{A.7}$$

Multiplying both sides of (A.6) by H and squaring:

$$X^2\left(\left(h_r\left(j\omega\right)+n_r\left(X\right)\right)^2+\left(h_i\left(j\omega\right)+n_i\left(X\right)\right)^2\right)$$
$$=\left(h_r^2\left(j\omega\right)+h_i^2\left(j\omega\right)\right)R^2$$

(A.8)

Jump resonance will first happen at $\dfrac{\partial R}{\partial X}=0$, a point where the output can change without a change in the input. So we differentiate with respect to X:

$$\left(h_r^2\left(j\omega\right)+h_i^2\left(j\omega\right)\right)R\frac{\partial R}{\partial X}=X\left\{\begin{array}{l}\left(h_r\left(j\omega\right)+n_r\left(X\right)\right)X\dfrac{\partial n_r\left(X\right)}{\partial X}\\[2mm]+\left(h_i\left(j\omega\right)+n_i\left(X\right)\right)X\dfrac{\partial n_i\left(X\right)}{\partial X}\\[2mm]+\left(h_r\left(j\omega\right)+n_r\left(X\right)\right)^2\\[2mm]+\left(h_i\left(j\omega\right)+n_i\left(X\right)\right)^2\end{array}\right\}$$

(A.9)

$$\left(h_r\left(j\omega\right)+n_r\left(X\right)+X\frac{\partial n_r\left(X\right)}{\partial X}\right)^2$$
$$=2\left(h_r\left(j\omega\right)+n_r\left(X\right)\right)X\frac{\partial n_r\left(X\right)}{\partial X}+\left(h_r\left(j\omega\right)+n_r\left(X\right)\right)^2+\left(X\frac{\partial n_r\left(X\right)}{\partial X}\right)^2$$

(A.10)

$$\left(h_r^2\left(j\omega\right)+h_i^2\left(j\omega\right)\right)R\frac{\partial R}{\partial X}=X\left\{\begin{array}{l}\left(h_r\left(j\omega\right)+n_r\left(X\right)+\dfrac{X}{2}\dfrac{\partial n_r\left(X\right)}{\partial X}\right)^2\\[3mm]+\left(h_i\left(j\omega\right)+n_i\left(X\right)+\dfrac{X}{2}\dfrac{\partial n_i\left(X\right)}{\partial X}\right)^2\\[3mm]-\left(\dfrac{X}{2}\dfrac{\partial n_r\left(X\right)}{\partial X}\right)^2-\left(\dfrac{X}{2}\dfrac{\partial n_i\left(X\right)}{\partial X}\right)^2\end{array}\right\}$$

(A.11)

$$v_r\left(X\right)=n_r\left(X\right)+X\frac{\partial n_r\left(X\right)}{\partial X}$$

(A.12)

$$v_i(X) = n_i(X) + X \frac{\partial n_i(X)}{\partial X} \tag{A.13}$$

$$\left(h_r^2(j\omega) + h_i^2(j\omega)\right) R \frac{\partial R}{\partial X} = X \left\{ \begin{array}{l} \left(h_r(j\omega) + \dfrac{n_r(X) + v_r(x)}{2}\right)^2 \\ + \left(h_i(j\omega) + \dfrac{n_i(X) + v_i(x)}{2}\right)^2 \\ - \dfrac{(n_r - v_r)^2 + (n_i - v_i)^2}{4} \end{array} \right\} \tag{A.14}$$

$$\left(h_r(j\omega) + \frac{n_r(X) + v_r(x)}{2}\right)^2 + \left(h_i(j\omega) + \frac{n_i(X) + v_i(x)}{2}\right)^2$$
$$= \frac{(n_r - v_r)^2 + (n_i - v_i)^2}{4} \tag{A.15}$$

When $\dfrac{\partial R}{\partial x} < 0$, jump resonance will still be possible and merely occur at a smaller value of V_{in},

Q, or ω at which $\dfrac{\partial R}{\partial x} = 0$. Thus, we replace the equality sign in A.15 with $\leq$ to get the criterion for jump resonance.

$$\left(h_r(j\omega) + \frac{n_r(X) + v_r(x)}{2}\right)^2 + \left(h_i(j\omega) + \frac{n_i(X) + v_i(x)}{2}\right)^2$$
$$\leq \frac{(n_r - v_r)^2 + (n_i - v_i)^2}{4} \tag{A.16}$$

111

References

[1]W. E. Scott, "An introduction to the analysis of non-linear closed cycle control systems," in *Automatic and Manual Control.* London: Butterworth, 1952, pp.249-261

[2] K. Horvat, M. Miskovic, and Ognjen Kuljaca, "Avoidance of nonlinear resonance jump in turbine governor positioning system using fuzzy controller." *IEEE Int. Conf. on Industrial Tech.* vol. 2 Dec. 2003 pp.881-886.

[3] R. Walk, R. J. Kakalec, and J. Rootenberg, "An analytic and computer study of jump phenomenon in ferroresonant regulators." *IEEE Trans. Magn.,* vol. 7, iss. 3, Sep. 1971 pp.574-577

[4] D. L. Hiser and R. L. Geiger, "Impact of OTA nonlinearities on the performance of continuous-time OTA-C bandpass filters." *IEEE Int. Symposium on Circuits and Systems.* vol. 2 May 1990 pp.1167-1170.

[5] R. S. Moni and K. Radhakrishna Rao, "Jump-phenomenon in Active-RC Filters." *IEEE Trans. Circuits Syst.* vol. 29 no.1 Jan. 1982 pp.54-55.

[6] S. Krawczyk, "Comment on 'Jump Phenomenon in Active-RC Filters.'" *IEEE Trans. Circuits Syst.* vol. 30 no.12 Dec. 1983 pp.908-909.

[7] A. Fukuma and M. Matsubara, "Jump Resonance Criteria of Nonlinear Control Systems," *IEEE Trans. Automat. Contr.,* vol. 11, iss. 4, Oct. 1966, pp. 699-706.

[8] C. D. Salthouse and R. Sarpeshkar, "A practical micropower programmable bandpass filter for use in bionic ears." *IEEE J. Solid-State Circuits,* vol.38 no.1 Jan 2003 pp.63-70.

[9]W. Germanovix and C. Toumazou, "Design of a Micropower Current-Mode Log-Domain Analog Cochlear Implant." *IEEE Trans. Circ. And Systems II Analog and Digital Signal Process.* vol. 47 no. 10 Oct. 2000 pp. 1023-1046

[10] R. Sarpeshkar, C. Salthouse, J.J. Sit, M. Baker, S. Zhak, T. Lu, L. Turicchia, and S. Balster, "An Ultra-Low-Power Programmable Analog Bionic Ear Processor", *IEEE Transactions on Biomedical Engineering*, Vol. 52, No. 4, pp. 711-727, April 2005.
[11] R. Sarpeshkar, M. Baker, C. Salthouse, J.J. Sit, L. Turicchia, and S. Zhak, "An Analog Bionic Ear Processor with Zero-Crossing Detection", *Proceedings of the IEEE International Solid State Circuits Conference (ISSCC)*, Paper 4.2, pp. 78—79, San Francisco, CA, February 6-10, 2005.

[12] J. S. Chang and Y. C. Tong, "A Micropower-Compatible Time-Multiplexed SC Speech Spectrum Analyzer Design." *IEEE J. Solid-State Circuits*, vol. 28 no.1 Jan 1993 pp. 40-48.

[13] B. Rodgers, S. Goenawan, M. Yunus, Y. Kaneko, and J. Yoshiike, "A 16-uA Interface Circuit for a Capacitive Flow Sensor." *IEEE J. Solid-State Circuits*, vol. 33 no. 12 Dec 1998 pp.2121-2133

[14] E. O Oladiran, E. Pisler, and S. Israelsson, "New Lightening Flash Counter and Calibration Circuit With Improved Discrimination of Cloud and Ground Discharges." *Science, Measurement and Technology, IEE Proceedings*, vol. 135 no.1 Jan 1988 pp. 22-28

[15] D. M. Binkley, J. M. Rochelle, B. K. Swann, L. G. Clonts, and R. N. Goble, "A Micropower CMOS, Direct-Conversion, VLF Receiver Chip for Magnetic-Field Wireless Applications." *IEEE J. Solid-State Circuits, vol. 33 no. 3* March 1998 pp. 344-358

[16] J. O. Pickles *An Introduction to the Physiology of Hearing*, New York: Academic Press, 1988

[17] M. A. Ruggero, S. S. Narayan, A. N. Temchin, and A. Recio, "Mechanical Bases Of Frequency Tuning And Neural Excitation At The Base Of The Cochlea: Comparison Of Basilar-Membrane Vibrations And Auditory-Nerve-Fiber Responses In Chinchilla." *Proceeding of the National Academy of Sciences*. vol. 97 no. 22 October 24, 2000 pp. 11744-11750.

[18] R. F. Lyon and C. Mead, "An Analog Electronic Cochlea." *IEEE Trans. On Acoustics, Speech, and Signal Processing*. vol. 36 no.7 July 1998 pp. 1119-1134.

[19] I. Grech, J. Micallef, and T. Vladimirova, "Silicon Cochlea And Its Adaptation To Spatial Localization" *IEE Proceedings of Circuits, Devices, and Systems* vol. 146 no. 2 April 1999 pp. 70-76

[20] R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A Low-Power Wide Dynamic-Range Analog VLSI Cochlea," *Analog Integrated Circuits and Signal Processing* vol. 16 no. 3 Aug. 1998 pp. 245-274

[21] T. Hirahara and T. Komakine, "A Computational Cochlear Nonlinear Preprocessing Model With Adaptive Q Circuits." *1989 Intl. Conf. on Acoustics, Speech, and Signal Processing* vol. 1 May 1989 pp. 496-499

[22] R. M. Kil, Y-I Kim, and G. H. Lee, "A Noise-Robust Front-End Based On Tree-Structured Filter-Bank For Speech Recognition." *Proc. Of IEEE-INNS-ENNS Intl. Joint Conf. on Neural Networks* vol. 6 July 2000 pp. 81-86

[23] E. Sanchez-Sinencio and J.Silva-Martinez, "CMOS transconductance amplifiers, architectures and active filters: A tutorial," *IEE Proc.-Circuits Devices Syst*, vol. 147, Feb. 2000. pp. 3-12

[24] C. Mead, *Analog VLSI and Neural Systems*, USA: Addison-Wesley Publishing Company, 1989 pp.67-69.

[25] L.C. Goldfarb, "On some nonlinear phenomena in regulatory systems" *Automat. Telemech.* Trans. In *Frequency Response*, R. Oldenburger, Ed. New York: Macmillan, 1956. pp. 349-383

[26] G. Palumbo and S. Pennisi, "High-Frequency Harmonic Distortion In Feedback Amplifiers: Analysis And Applications," *IEEE Trans. On Circuits and Syst. I.* vol. 50 no. 3 March 2003 pp.328-340

[27] S. M. Zhak, M. W. Baker, and R. Sarpeshkar, "A low-power wide dynamic range envelope detector." *IEEE J. Solid-State Circuits.* vol. 38 iss. 10 Oct. 2003 pp.1750-1753.

[28] Turicchia, L., and R. Sarpeshkar, "A Bio-Inspired Companding Strategy for Spectral Enhancement", *IEEE Transactions on Speech and Audio Processing*, Vol. 13, No. 2, March 2005.

[29] R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A low-power wide-linear-range transconductance amplifier," *Analog Integrated Circuits Signal Process.*, vol. 13, no.1/2, May/June 1997.

114

[30]G. Groenewold, "Optimal dynamic range integrators." *IEEE Trans. Circuits Syst. vol. 39 iss. 8 Aug. 1992 pp.614-627.*

[31]C. Hayahsi, *Nonlinear Oscillations in Physical Systems*, New York: McGraw-Hill Book Company, 1964 pp. 209-237

Chapter 4 : Companding Speech Processor

As we speak, we change the shape of our vocal tract and the method of excitation to create the sounds of speech. For vowels and voiced consonants, the vocal tract is excited at one end by the periodic movements of the vocal folds. From the vocal folds to the lips, the vocal tract may be modeled as a tube that is closed at one end and open at the other. It generates expected resonances with characteristic wavelengths near twice the tract length, three halves the length, five halves the length, et cetera. We change the shape of this tube by moving our soft pallet, tongue, lips, and jaw in order to move the resonant peaks. As we listen, we identify the sounds of speech by detecting the frequency peaks [1].

Hearing loss makes this process more difficult in a few ways. First, the frequency resolution is decreased even within the remaining hearing range by the loss of hair cells. Then, the problem is made worse by the dynamic range compression used by hearing aids and cochlear implants. Hearing loss raises the threshold below which the patient cannot hear, but the maximum loudness level above which sounds are uncomfortable does not increase. In order for patients to hear a normal range of sound intensities, the hearing device must amplify quiet signals more than large signals [2]. This decreases the intensity differences between the peaks and valleys in the frequency content of speech, referred to as spectral contrast [3].

Speech intelligibility always decreases with noise, but for hearing aid and cochlear implant users the decrease occurs at lower noise levels. Turicchia and Sarpeshkar have proposed a companding algorithm that improves performance in this moderate-noise regime [2]. The companding algorithm is a highly tunable system that exhibits two-tone suppression, a behavior found in healthy ears where the amplitude of smaller tones is decreased in the presence of larger tones. It is one of the techniques used by healthy ears to improve spectral contrast.

The companding algorithm is performing very well in preliminary tests. So far it has been tested by researchers at the Massachusetts Institute of Technology, the University of California at San Diego, and the University of Texas at Dallas. Each group has shown statistically significant improvements in speech recognition scores in the moderate noise regime [4].

In order for the companding algorithm to help hearing aid and cochlear implant users, it must be implemented in a technology compatible with those applications. It must be small and low power while preserving as much of the flexibility and performance of the computer based implementations. This paper presents the micropower implementation of the companding algorithm in analog CMOS that can run for 4 months on a standard hearing aid battery. I first present the companding algorithm, section 1. Then I present the analog building blocks used in this implementation, section 2. In section 3, I present the system design. I report the experimental results for the single channel in section 4 and the system performance in section 5.

Section 1. Companding Algorithm

Turicchia and Sarpeshkar proposed a companding algorithm for performing compression and spectral contrast enhancement for cochlear implants, hearing aids, and speech recognition systems [2]. The algorithm processes sound in parallel channels as shown in Fig. 4.1. It performs two-tone suppression in a two step procedure. The input signal, X_0, is connected to the input of all of the compression channels. Then a broad filter selects a broad range of frequencies, X_1, which are compressed by amplifying small signals and attenuating large signals, X_2. Then the expansion half channel selects a narrower band of frequencies, X_3, and expands the dynamic range by attenuating small signals and amplifying large signals, X_4.
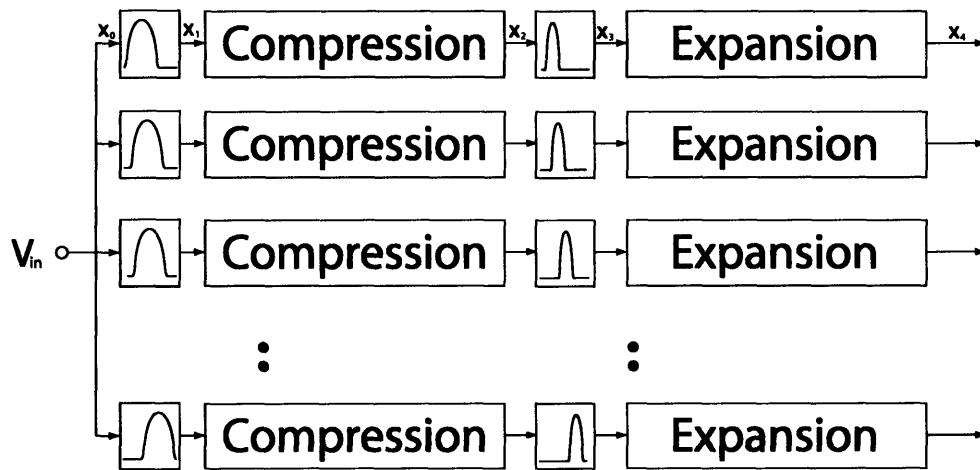
Figure 4.1: Block diagram demonstrating the parallel channels. Sound enters at V_{in} and is first filtered by broad input filters. The filtered signals are then compressed. A narrow bandpass filter selects a subset of these compressed frequencies. The narrowband signal is then expanded.

118

The mechanism of two-tone suppression is revealed by considering a few cases. The signals at each of the nodes labeled in Fig. 4.1 are shown in the frequency domain in Fig. 4.2. For each of these cases, the system is biased such that any single tone has the same amplitude at the output as at the input. The tone passes through multiple compression channels because of the broad input filters, but the narrow-band expansion filter limits the majority of the signal energy to only one or two channels. The signals in Fig. 4.2 are for the channel where both filters are centered at the tone frequency, as shown by the grey filter curves in each block. In case A, a small tone is placed at the input of the system. This tone passes through the broadband filter to X_1. Then the tone is amplified by the compressor to form X_2. The narrowband filter creates X_3 with the same amplitude as X_2. Finally, the expander attenuates the signal to create X_4 with the same amplitude as X_0. In case B, a large signal is attenuated in the compression stage and amplified back to its original amplitude in the expansion stage.
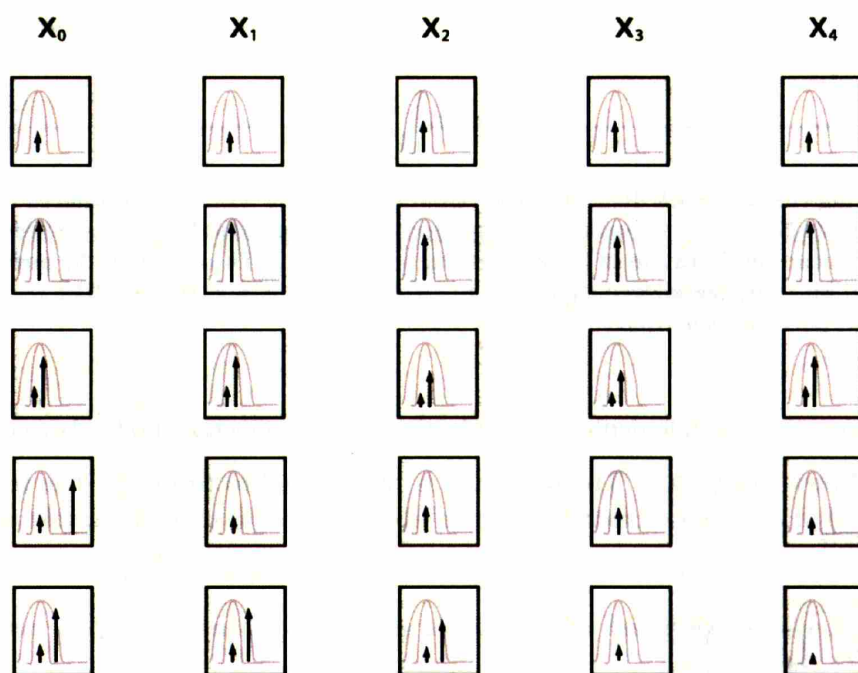
Figure 4.2: Companding Test Cases

Three different cases can be formed by summing two sinusoids, which I will refer to as the large tone and the small tone, to form the input signal. In case C, both tones fall within the same narrow band filter. In this case, the large tone leads to attenuation in the compressor and amplification in the expander so the system has a unity gain. In case D, the tones do not fall

119

within the same broad filters. In this case, the tones are processed independently with different channels. Companding is demonstrated in case E, when the tones fall within the same broadband filter, but different narrowband filters. As in all of the cases, the signal is compressed by multiple channels because of the broad compression filter. The large tone causes all of these channels to attenuate the signal. In this case, the expansion filter in one of these channels selects the large tone; I will call this the large channel. The expansion filter in the small channel, shown in Fig. 4.2, selects the small tone. In the large channel, the signal is again amplified, creating unity gain. In the small channel, only the small tone is present so it is attenuated. In this channel, the signal has been attenuated twice leading to a decrease in the amplitude of small signals in the proximity of large signals, i.e. two-tone suppression.
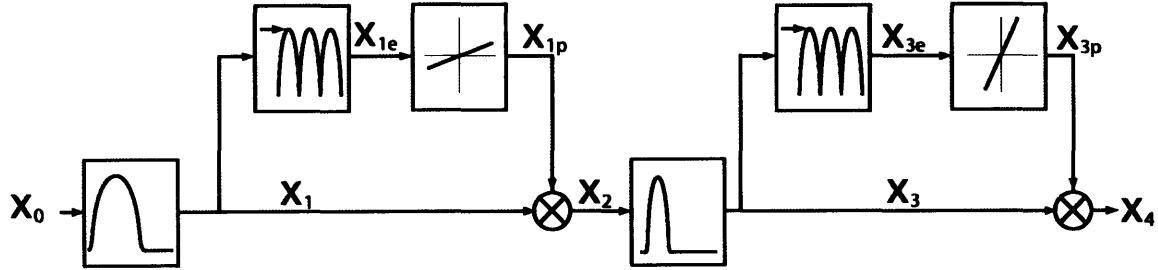
Figure 4.3: Companding channel block diagram. The signal enters at X_0. The signal is filtered by the broadband filter to create X_1. The envelope detector creates the envelope signal, X_{1e}. The power law circuit then applies a linear amplification in the log domain to create X_{1p}. The output of the power law circuit, X_{1p}, is multiplied by the filtered signal, X_1, to form a compressed signal, X_2. The expansion half channel performs simarily.

The companding channels are implemented using filters, envelope detectors, power law circuits, and multipliers as shown in Fig. 4.3. The input, X_0, is first filtered to form X_1. The transfer function of this filter is

$$F(s) = \left(\frac{2\left(\frac{\tau}{q_f} \right)s}{\tau^2 s^2 + 2\left(\frac{\tau}{q_f} \right)s + 1} \right)^2$$

(4.1)

The filter is defined by a time constant, τ, and bandwidth factor, q_f, as proposed in the companding definition [2]. It is important to note that this q_f is not the same as the quality factor defined as the center frequency divided by the 3dB bandwidth.

120

The compression is performed by the envelope detector, power law block, and multiplier. The envelope detector creates an envelop signal, X_{1e}, by measuring the amplitude of X_1. This envelope signal is then raised to a power in the power law block.

$$X_{1p} = \left(X_{1e}\right)^{n_1-1} \tag{4.2}$$

The companding index, n_1, sets the amount of two-tone suppression. The filtered input, X_1, is then multiplied by this power law signal, X_{1p} to form the compressed signal, X_2. Its amplitude is

$$|X_2| = |X_1| \cdot X_{1p} = |X_1| \cdot |X_1|^{n_1-1} = |X_1|^{n_1} \tag{4.3}$$

The expansion half channel performs the same functions except that the power law circuit produces an output

$$X_{3p} = \left(X_{3e}\right)^{\frac{n_2-n_1}{n_1}} \tag{4.4}$$

Where n_2 is the compression index, the parameter which sets the net compression for the channel. When the compression index is equal to 1, there is no net compression or expansion

$$|X_4| = |X_3| \cdot X_{pe}^{\frac{1-n_1}{n_1}} = |X_3|^{\frac{1}{n_1}} = |X_2|^{\frac{1}{n_1}} = |X_1|^{\frac{n_1}{n_1}} = |X_1| \tag{4.5}$$

The analog implementation of this system must allow all of these parameters to be adjusted, but it is helpful to start with the values used in the original paper. In each channel, the filters had the same center frequency varied from 250Hz to 4kHz. The broad filters have a q of 2.8. The narrow filter poles have a q of 4.5. The companding index was 0.25 and the compression index was 1.

For micropower design, it is also necessary to specify the dynamic range of the system. In filters in particular it has been shown that power scales with dynamic range [4-5]. A healthy human can hear sounds over a range of greater than 100dB [7]. But, the variation within a speech utterance is only about 30dB [1]. An automatic gain control, such as the one presented

in [8], may be used to compress real world speech into the 40dB used in a micropower companding system.

Section 2. Analog Building Blocks

I have implemented this system in analog CMOS for micropower operation. Both filters are realized as Gm-C filters as described in subsection A. Subsection B details the combination of a rectifier and a peak detector that performs the envelope detection. Subsection C presents the power law circuit that calculates the require gain for compression and expansion. Finally, subsection D presents the variable gain amplifier that performs the multiplication.

A: Filters

The filters were implemented using differential transconductors and capacitors. The basic topology of the transconductors is shown in Fig. 4.4. The bias current is divided between the two current arms by M_1 and M_2. The current mirror subtracts the negative current leg from the positive leg to create I_{out}. The transconductance is

$$G_m = \frac{I_{bias}\kappa}{2\phi_t} \tag{4.6}$$

where κ is the subthreshold exponential constant for the PMOS transistors, typically around 0.7, and ϕ_t is the thermal voltage, approximately 25mV. And circuit is approximately linear for

$$|V_+ - V_-| \leq V_l = \frac{2\phi_t}{\kappa} \tag{4.7}$$

Approximately 75mV [9]

122
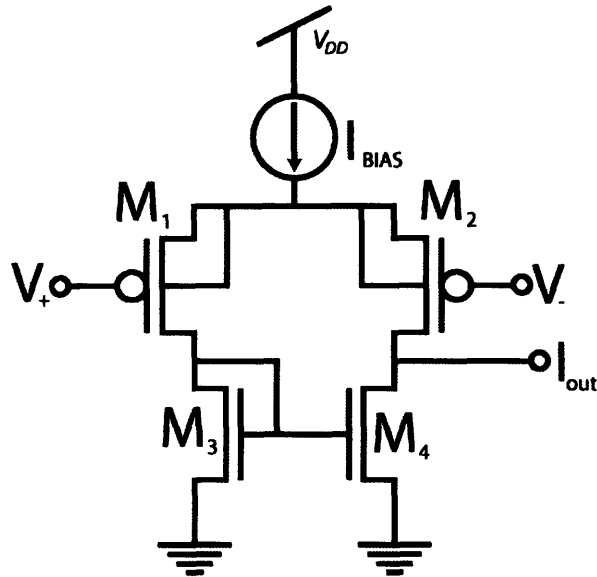
Figure 4.4: Basic transconductor topology. The output current is proportional to the differential input voltage.
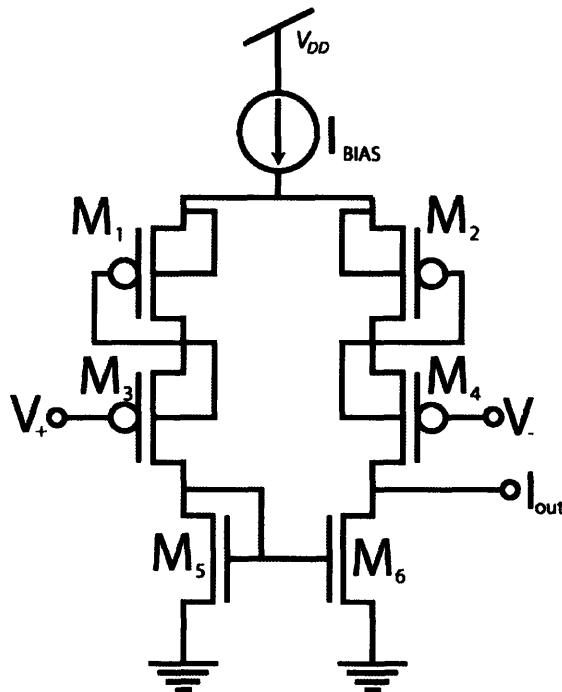
Figure 4.5: Degenerated transconductor topology. This transconductor is shown to lead to more power efficient filters.

Source degeneration, as shown in Fig. 4.5, doubles the linear range for the transconductors. It also halves the transonductance. For this application, the additional linear range is a minor improvement, but the power requirements for a given dynamic range are important. The

power requirements for the degenerated transconductor can be compared to the simple transconductor by calculating the input referred noise for each of the two structures.

In subthreshold operation, the noise in these transistors is well modeled as shot noise sources in parallel with each transistor with amplitudes given by

$$I_{noise}^2 = 2qI_D \tag{4.8}$$

where q is the charge of an electron and I_D is the drain current of the transistor. The noise sources are shown for the degenerated transconductors in Fig. 4.6 [10].
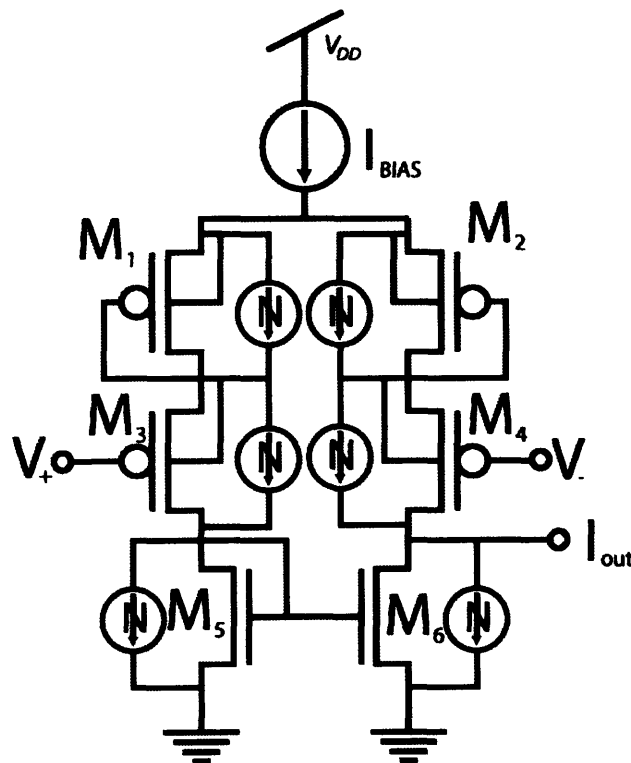
Figure 4.6: Degenerated transconductor with noise sources. The noise introduced by each transistor can be modeled as a noise current source in parallel with the transistor. The total output current noise can be calculated by adding the power contributions from each noise source.

Because each noise source is independent, the total noise at the output is calculated by summing the squared contributions of each noise source. In this circuit, the degeneration causes the transfer function from each of the input and degeneration transistors to be ½ leading to the equation for the output noise

$$I_{noise}^2 = 4\left(2q\frac{I_{bias}}{2}\left(\frac{1}{2}\right)^2\right) + 2\left(2q\frac{I_{bias}}{2}\right) = 3qI_{bias} \tag{4.9}$$

This noise can be referred to the input by dividing by the transconductance squared

$$V_{noise}^2 = 3qI_{bias}\left(\frac{4\phi_t}{I_{bias}\kappa}\right)^2 = \frac{48q\phi_t^2}{I_{bias}\kappa^2} \tag{4.10}$$

The dynamic range of a filter is defined as the difference between the largest signal the system can filter and the smallest signal that can be detected at the output. The largest signal is proportional to the linear range, V_l. The smallest signal can be defined by the total noise at the output. With any transconductors topology, the noise can be reduced by using a larger capacitor which averages over more charge particles. The larger capacitor increases the required power consumption by requiring a larger transconductance. Thus, the figure of merit must also include a term for the transconductance. The figure of merit for transconductors is given by

$$FOM = \frac{GV_l}{\sqrt{V_{noise}^2}} = \frac{I_{bias}}{\frac{4\phi_t}{\kappa}\sqrt{\frac{3q}{I_{bias}}}} = \frac{\kappa I_{bias}^{\frac{3}{2}}}{4\phi_t\sqrt{3q}} \tag{4.11}$$

For the simple transconductors of Fig. 4.4, the output noise is

$$I_{noise}^2 = 2q\frac{I_{bias}}{2}4 = 4qI_{bias} \tag{4.12}$$

The input referred noise is

$$V_{noise}^2 = 4qI_{bias}\left(\frac{2\phi_t}{I_{bias}\kappa}\right)^2 = \frac{16q\phi_t^2}{I_{bias}\kappa^2} \tag{4.13}$$

The figure of merit for the simple transconductors is

125

$$FOM = \frac{GV_I}{\sqrt{V_{noise}^2}} = \frac{I_{bias}}{\sqrt{\dfrac{16q\phi_t^2}{I_{bias}\kappa^2}}} = \frac{\kappa I_{bias}^{\frac{3}{2}}}{4\phi_t\sqrt{q}}$$ (4.14)

This figure of merit demonstrates that the degenerated transconductor can be used to build a filter with the same dynamic range as the simple transconductor, but lower power consumption.
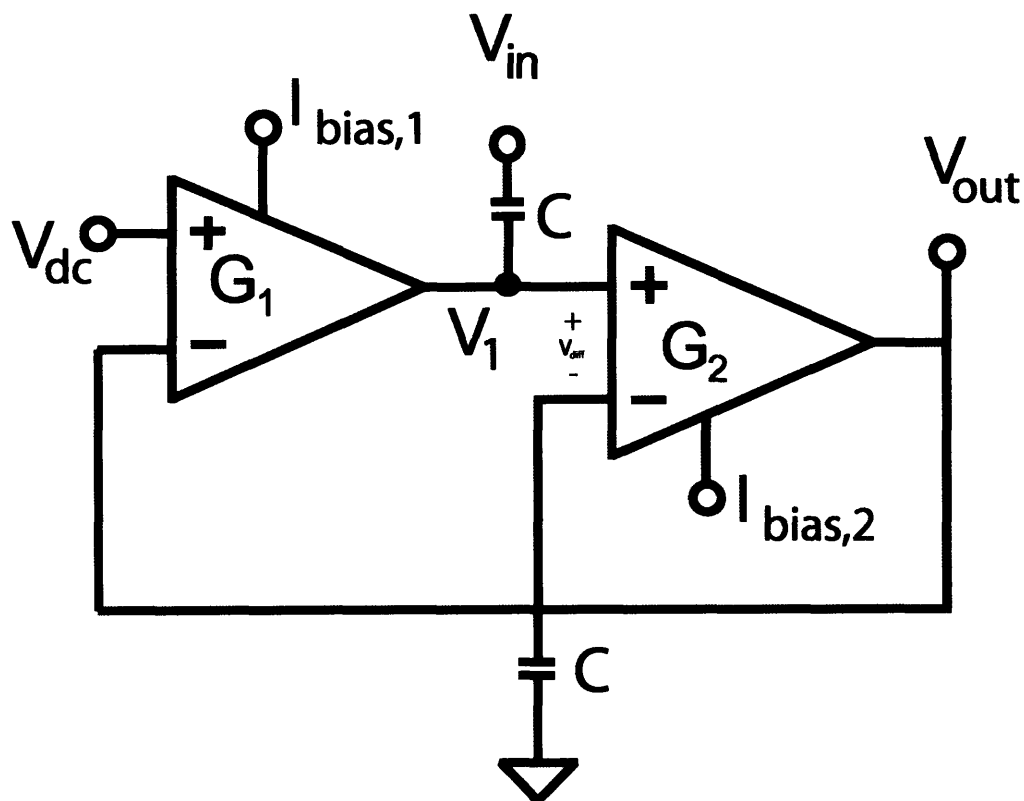
Figure 4.7: Basic filter topology. Two transconductors are combined with two capacitors to form this bandpass filter.

The topology for such a bandpass filter is shown in Fig. 4.7. The input is capacitively coupled to V_1. This signal is then lowpass filtered by G_2 acting as an RC filter. A low frequency path through G_1 removes low frequency components from V_1, completing the bandpass. The loop can be analyzed in more detail by creating a block diagram for the circuit as in Fig. 4.8.
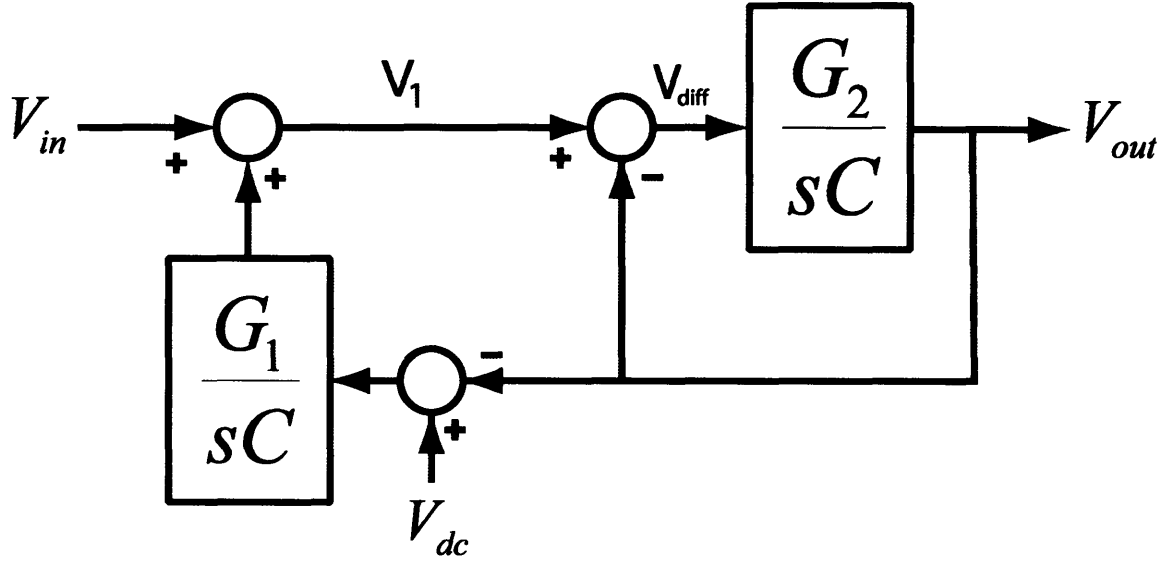
Figure 4.8: Basic filter block diagram. The feedback found in the filter in Fig. 6 can be seen clearly in this block diagram.

The transfer function from the input to the output of the filter is

$$H(s) = \frac{G_2 sC}{G_2 G_1 + sCG_2 + s^2 C^2}$$

(4.15)

This can be compared to the first half of the desired transfer function

$$F(s)^{\frac{1}{2}} = \frac{2\left(\frac{\tau}{q_f}\right)s}{\tau^2 s^2 + 2\left(\frac{\tau}{q_f}\right)s + 1}$$

(4.16)

The values of the center frequency and q can be calculated as a function of the bias currents.

$$G_1 = \frac{I_1 \kappa}{4\phi_t}$$

(4.17)

$$G_2 = \frac{I_2 \kappa}{4\phi_t} \tag{4.18}$$

$$f_{filt} = \frac{1}{2\pi} \sqrt{\frac{G_1 G_2}{C^2}} = \frac{1}{2\pi} \frac{\kappa}{4\phi_t C} \sqrt{I_1 I_2} \tag{4.19}$$

$$q_f = 2\sqrt{\frac{G_1}{G_2}} = 2\sqrt{\frac{I_1}{I_2}} \tag{4.20}$$
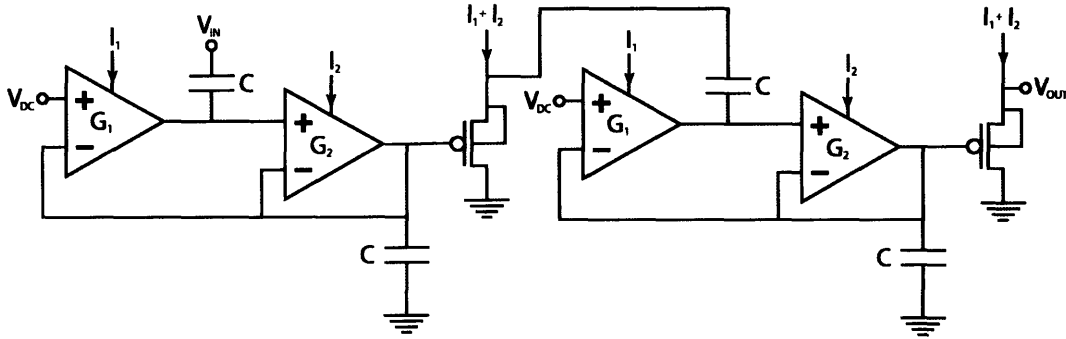
Figure 4.9: Cascaded filters. Two filters are cascaded to create second-order roll-off. A buffer is inserted between the filters to prevent loading.

Two filters can be cascaded as shown in Fig. 4.9 to form the filter specified in the companding algorithm with second order roll-off and greater out-of-band suppression. A buffer is used to couple the filters without loading. Because the center frequency is determined by the geometric mean of the two currents, the buffer can be guaranteed to have a greater bandwidth than the filter if it is biased with the sum of the currents with an appropriate scale factor. The frequency of this parasitic buffer pole is then given by

$$f_{parasitic} = \frac{G_{fol}}{2\pi C} = \frac{(I_1 + I_2)\kappa}{2\pi C (\phi_t)} \tag{4.21}$$

This pole frequency can be compared to the center frequency of the filter by writing the currents in terms of filter parameters.

$$I_1 = f_{filt} q_f \frac{4\pi \phi_t C}{\kappa} \tag{4.22}$$

128

$$I_2 = \frac{f_{filt}}{q_f} \frac{16\pi\phi_t C}{\kappa} \tag{4.23}$$

Then the parasitic pole location can be written in terms of filter parameters to demonstrate that the parasitic pole is at least a factor of 8 higher than the center of the filter if the buffer is biased with unity scale factor.

$$f_{parasitic} = \frac{\left(f_{filt} q_f \frac{4\pi\phi_t C}{\kappa} + \frac{f_{filt}}{q_f} \frac{16\pi\phi_t C}{\kappa} \right)\kappa}{2\pi C(\phi_t)} = 4f_{filt}\left(\frac{q_f}{2} + \frac{2}{q_f} \right) \tag{4.24}$$

$$f_{parasitic} \geq 8f_{filt} \tag{4.25}$$

The maximum signal for which the filter is linear is set by the smallest signal that creates an input to a transconductors that is larger than the transconductor's linear range. This can be found by calculating the transfer function from the differential input of each transconductor to the output of the filter. The transfer function from G_1 to the output is

$$\frac{V_{out}}{V_{G_1}}(s) = \frac{1}{1 + s2\frac{\tau}{q_f} + s^2\tau^2} \tag{4.26}$$

The maximum value of this transfer function is approximately equal to $q_f/2$. The transfer function from G_2 to the output is

$$\frac{V_{out}}{V_{G_2}}(s) = \frac{2s\frac{\tau}{q_f}}{1 + 2s\frac{\tau}{q_f} + s^2\tau^2} \tag{4.27}$$

The maximum value of this transfer function is 1, the same as the bandpass filter. So, the maximum signal at the output of the filter is approximately the linear range of the transconductors, 100mVrms, divided by half the q_f of the filter.

The noise of filter can be calculated based on the noise from each transconductors. First, the noise of the basic filter block shown in Fig. 4.7 is calculated

$$V^2_{noise,filtout} = \frac{48q\phi_t^2}{I_1\kappa^2}\left(\left|\frac{G_1G_2}{G_1G_2 + sCG_2 + s^2C^2}\right|\right)^2 + \frac{48q\phi_t^2}{I_2\kappa^2}\left(\left|\frac{sCG_2}{G_1G_2 + sCG_2 + s^2C^2}\right|\right)^2$$

$$= \frac{48q\phi_t^2}{I_1I_2\kappa^2}\frac{I_1\left(\left|sCG_2\right|\right)^2 + I_2\left(G_1G_2\right)^2}{\left(G_1G_2 + sCG_2 + s^2C^2\right)^2}$$

(4.28)

Then the noise of the buffer is calculated first as an output current noise

$$I^2_{noise} = 2\left(2q\left(I_1 + I_2\right)\right)$$

(4.29)

Then the output current noise is referred to the input of the buffer

$$V^2_{noise,buf} = \frac{4q\left(I_1 + I_2\right)}{G^2} = \frac{4q\left(I_1 + I_2\right)\phi_t^2}{\kappa^2\left(I_1 + I_2\right)^2} = \frac{4q\phi_t^2}{\kappa^2\left(I_1 + I_2\right)}$$

(4.30)

The noise at the output of two filters and two buffers is:

$$V^2_{noise,total} = \left(\frac{48q\phi_t^2}{I_1I_2\kappa^2}\frac{I_1\left(\left|sCG_2\right|\right)^2 + I_2\left(G_1G_2\right)^2}{\left(G_1G_2 + sCG_2 + s^2C^2\right)^2} + \frac{4q\phi_t^2}{\kappa^2\left(I_1 + I_2\right)}\right)$$

$$\cdot\left(1 + \frac{\left(\left|sCG_2\right|\right)^2}{\left(G_1G_2 + sCG_2 + s^2C^2\right)^2}\right)$$

(4.31)

Integrating this noise over the frequencies of interest gives a total noise for the filter. With 2pF of capacitance, this filter has 189μVrms of noise for q_f=2. The dynamic range of the filter is nearly 54dB for q_f =2 going down to 51dB at the desired q_f =2.8 because of decreases in the linear range of the filter.

The bias currents can be calculated for the filters at 1kHz using Eq. 4.17 and Eq. 4.18. First, the broad filter with a Q of 2.8 has bias currents of

130

$$I_1 = f_{filt} q_f \frac{4\pi \phi_t C}{\kappa} = 10^3 (2.8) \frac{4\pi (0.025) 2 \cdot 10^{-12}}{0.7} = 2.5nA \tag{4.32}$$

$$I_2 = \frac{f_{filt}}{q_f} \frac{16\pi \phi_t C}{\kappa} = 1.2nA \tag{4.33}$$

The current consumption is

$$I_{total} = 2\left(I_1 + I_2 + \left(I_1 + I_2\right)\right) = 4\left(I_1 + I_2\right) = 4\left(3.7nA\right) = 14.8nA \tag{4.34}$$

In addition the translinear circuit that creates these currents from tau and Q currents consumes six times the tau current, where tau is the geometric mean of I1 and I2. 1.7nA in this case leading to 10.2 nA in tau and Q plus six times the Q current (6*1.4nA)=8.4nA so the total current is 33.4nA.

The total power consumed from a 2.8V supply for the broad filter at 1kHz is

$$P = IV = 33.4nA\left(2.8V\right) = 94nW \tag{4.35}$$

The same analysis for the narrow filter with a q_f of 4.5 gives a power consumption of 125nW. The total power consumed in filters for a sixteen channel system with filters spaced between 250Hz and 4kHz is

$$Power = \sum (94 + 125) nW \frac{(freq)}{1000} = 4.9\mu W \tag{4.36}$$

B: Envelope Detectors

In order to perform compression and expansion, the gain of each stage must be varied with the signal amplitude. This is done by first measuring the signal amplitude with an envelope detector (ED). The detectors used here are based on those developed by Zhak et. al. for use in cochlear implants[11-12]. The circuit topology shown in Fig. 4.10 differs from that presented earlier only in the use of the full wave output.

131
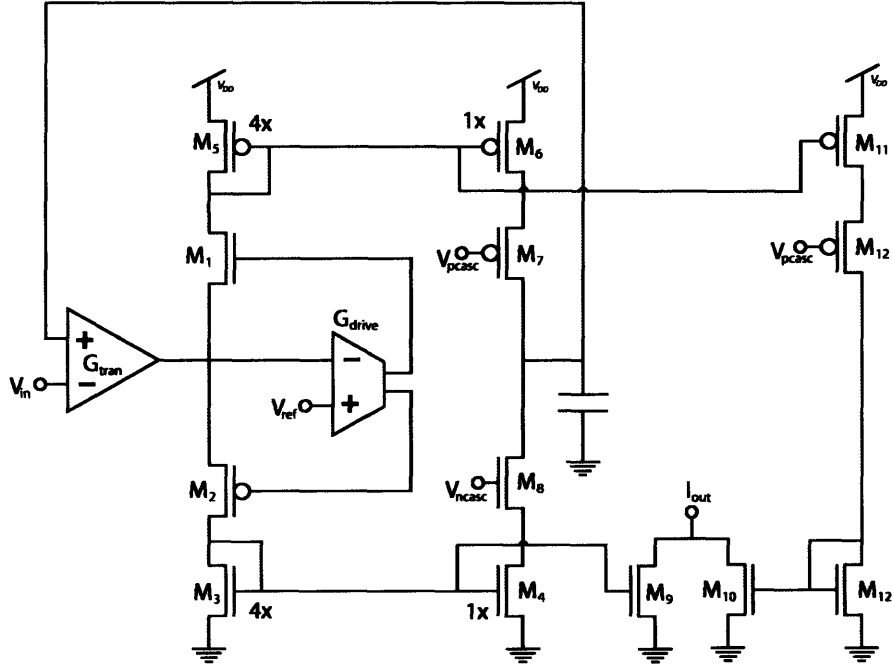
Figure 4.10: Envelope detector topology. This topology was developed by Zhak in [10].

In the operating frequencies, the input transconductor, G_{tran}, produces a current proportional to the input voltage. This current is either sourced through M_1 or sunk through M_2 depending on the sign. The current mirrors then create a rectified output current at I_{out}. The remaining components improve the performance. The capacitor is used as part of a lowpass filter from V_{in} to the non-inverting input of G_{tran}. This allows the system to reject DC offsets. Despite the use of the non-inverting input to the transconductor, this loop is stable negative feedback because of the sign flip introduced by the current mirrors. The drive transconductor decreases the minimum signal and increases the maximum frequency by actively driving the capacitance required to switch M_1 and M_2. As in the original ED design, the current mirrors M_3/M_4 and M_5/M_6 are 4 to 1 to reduce the capacitor size for a given cut-off frequency.

Changes were made in implementing the circuit building blocks of Fig. 4.10 to minimize power consumption and area for the 40dB dynamic range goal. First, G_{tran} was implemented with the same type of source degenerated transconductors as was used in the filters, Fig. 4.5, but with the addition of output current mirrors for wide output swing. Because the linear range is limited to the linear range of this transconductors by the filter, the extra linear range of a Wide-Linear-Range transconductors would have been wasted [5]. The output current mirrors were

132

set to a gain of 1, as opposed to N=5 in [11], in order to minimize the capacitor size and power consumption. The combination of unity current mirrors and smaller bias currents for a smaller dynamic range allowed the capacitance to be reduced from 180pF to 10pF. Finally, the drive transconductor, G_{drive}, was redesigned from the 1.5 μm design in [11] to provide the appropriate level shifting for the 0.5 μ process. The new topology is presented in Fig. 4.11.

Figure 4.11: Drive transconductor topology. M₉ provides shifted outputs.

The power consumption of the ED is going to be limited by the currents required to overcome parasitic capacitance. Decreasing the current through the input transconductor will lower all of the currents in the system, increasing the input referred dead zone. Conversely, increasing the current will increase the highpass cut-off of the circuit. This sets a maximum current for the input transconductor such that all frequencies of interest are above the cutoff.

$$\frac{G_{tran}}{4C} \leq 2\pi 250 \tag{4.37}$$

$$G_{tran} \leq (2\pi 1000)(10^{-11}F) \approx 62.8nS \tag{4.38}$$

$$Ibias \leq (6.28 \cdot 10^{-8})(150mV) \approx 10nA \tag{4.39}$$

133

The total power consumption also includes a factor of two from the output legs of the transconductors. Then, assuming that the drive transconductor operates at similar current levels, the power consumption of the envelope detectors on the chip will be the

$$Power_{ED} = 2.8V\left(16channels\right)\left(2halfs\right)\left(2(10nA)+2(10nA)\right) = 3.6\mu W \tag{4.40}$$

After rectification, the peak detector measures the amplitude using a current mode filter with asymmetric attack and release as shown in Fig. 4.12[11]. When I_{in} is increasing, the behavior of the circuit is governed by dynamic translinear circuit formed by M_6, M_7, M_8, M_9, and C_1. The frequency response in this regime can be calculated starting with the translinear equation

$$I_6 I_8 = I_7 I_9 \tag{4.41}$$

Substituting in values from this circuit

$$I_{in} I_{atk} = I_7 I_{out} \tag{4.42}$$

$$I_7 = I_{atk} + C_1 \frac{\partial V_{C_1}}{\partial t} \tag{4.43}$$

$$I_{out} = I_s e^{\frac{\kappa\left(V_{C_1}+V_x-V_{ref}\right)}{\phi_t}} \tag{4.44}$$

$$\frac{\partial I_{out}}{\partial t} = \frac{\kappa I_{out}}{\phi_t}\left(\frac{\partial V_{C_1}}{\partial t}\right) \tag{4.45}$$

$$I_7 = I_{atk} + \frac{C_1 \phi_t}{\kappa I_{out}}\frac{\partial I_{out}}{\partial t} \tag{4.46}$$

$$I_{in} I_{atk} = I_{atk} I_{out} + \frac{C_1 \phi_t}{\kappa}\frac{\partial I_{out}}{\partial t} \tag{4.47}$$

134

$$\frac{I_{out}}{I_{in}}(s) = \frac{I_{atk}}{I_{atk} + s\dfrac{C_1\phi_t}{\kappa}} = \frac{1}{1 + s\dfrac{C_1\phi_t}{I_{atk}\kappa}}$$

(4.48)

When I_{in} is decreasing, the translinear loop is no longer active. Instead, the output current decreases as I_{rel} increases the gate voltage on M_9. Because the gate voltage increases linearly, the output current decreases exponentially with a time constant of

$$\tau_r = \frac{C_2\phi_t}{I_{rel}\kappa}$$

(4.49)
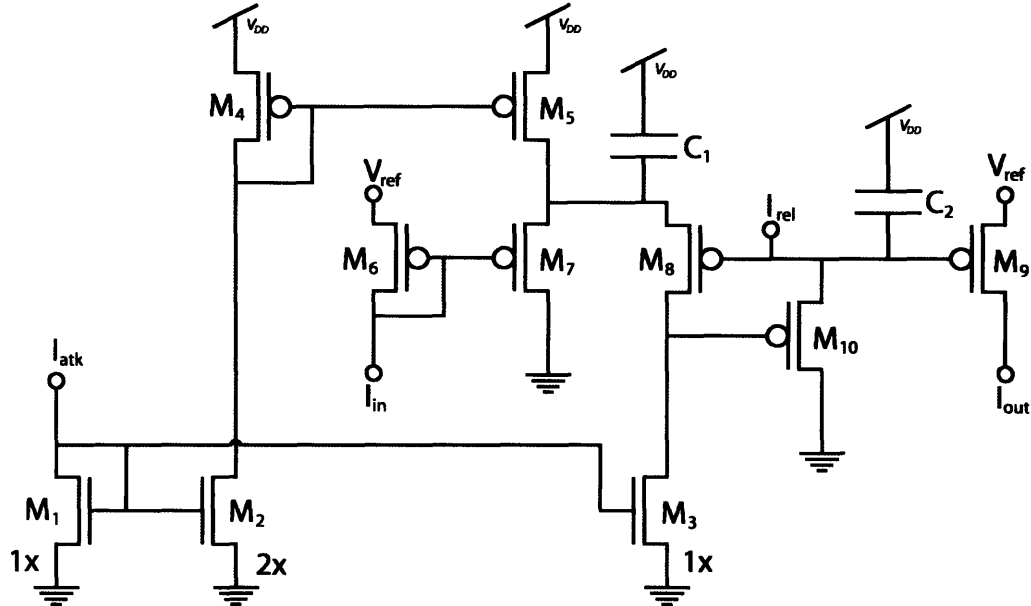
Figure 4.12: Peak detector circuit. This circuit was developed by Zhak et. al [10]. Currents are filtered with an attack time constant set by I_{atk} and a release time constant set by I_{rel}.

As with the rectifier, capacitances in this circuit were reduced from those reported earlier. C_1 was reduced from 30pF to 5pF. C_2 was reduced from 10pF to 5pF. The attack and release currents are

$$I_{atk} = \frac{C_1\phi_t}{\kappa\tau_{atk}} = \frac{5\cdot10^{-12}(0.025)}{0.7(0.001)} \approx 200pA$$

(4.50)

$$I_{rel} = \frac{C_2 \phi_t}{\kappa \tau_{rel}} = \frac{5 \cdot 10^{-12}(0.025)}{0.7(0.040)} \approx 5pA \qquad (4.51)$$

The peak detectors are predicted to consume

$$Power_{pd} = 16channels(2halves)(4 \cdot 200pA + 5pA) = 26nW \qquad (4.52)$$

C: Power Law Circuits

The power law circuits, presented in [8] and shown in Fig. 4.13, raise the envelope current to a current selectable power. The circuits operate by first taking the natural logarithm of the current using M_1. The resulting voltage is then compared to M_2 to generate a current with G_{in}. This current is converted back to a voltage by G_{out}, applying a voltage gain of G_{out}/G_{in}. This voltage is exponentiated in M_4 to create the output current. This operation can be written mathematically as

$$I_{out} = I_{ref,out} e^{\frac{G_{out}}{G_{in}}(\ln I_{in} - \ln I_{ref,in})} \qquad (4.53)$$
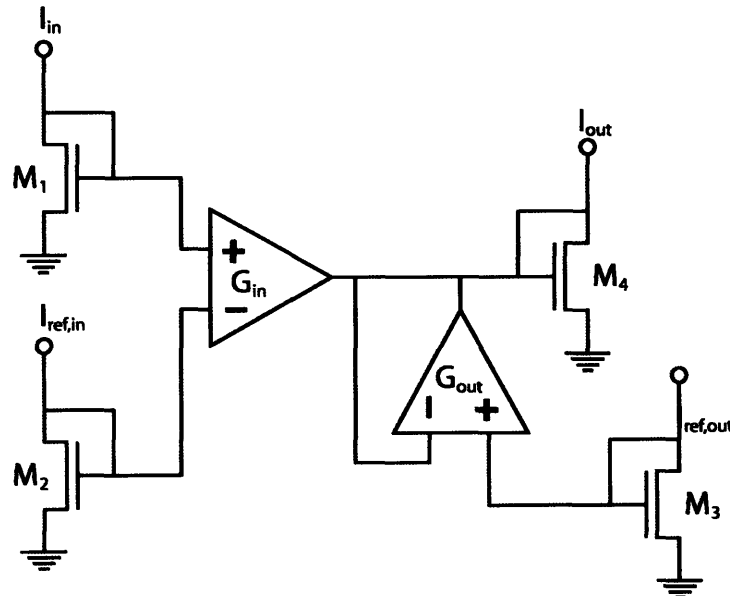

Figure 4.13: Power law circuit. This circuit was developed by Baker et. al. in [8]. The logarithm of the input current is taken by M_1. A current is generated by G_{in} comparing log(I_{in}) to log($I_{ref,in}$). This current is converted to a voltage by G_{out}. The output current is then related to the input current by a power law.

136

The input transconductor, G_1, must be linear for input currents variations from 1 to 100 for the 40dB required operation. The total range can be calculated using the subthreshold MOSFET equation

$$I_{in} = I_s e^{\frac{\kappa V_g}{\phi_t}} \tag{4.54}$$

$$100 I_{in} = 100 I_s e^{\frac{\kappa V_g}{\phi_t}} = I_s e^{\left(\frac{\kappa V_g}{\phi_t} + \ln 100\right)} = I_s e^{\frac{\kappa}{\phi_t}\left(V_g + \frac{\phi_t \ln 100}{\kappa}\right)} \tag{4.55}$$

The total swing in input voltage is

$$\frac{\phi_t \ln 100}{\kappa} \approx 165mV \tag{4.56}$$

By placing $I_{ref,in}$ in the middle of this range the linear range requirement can be brought down to 83mV, slightly larger than the simple transconductor. So a source degenerated transconductor is used here as well.

The power requirements for the power law circuit are set by the currents required to drive the variable gain amplifier. The other currents may be small, since the envelope signal has a limited bandwidth.

D: Variable Gain Amplifier

The current derived from the signal envelope and raised to a programmable power is used to control a variable gain amplifier (VGA) [8]. As in the power law cell, the input transconductor creates a current which is converted back to a voltage by the output transconductor as shown in Fig. 4.14.
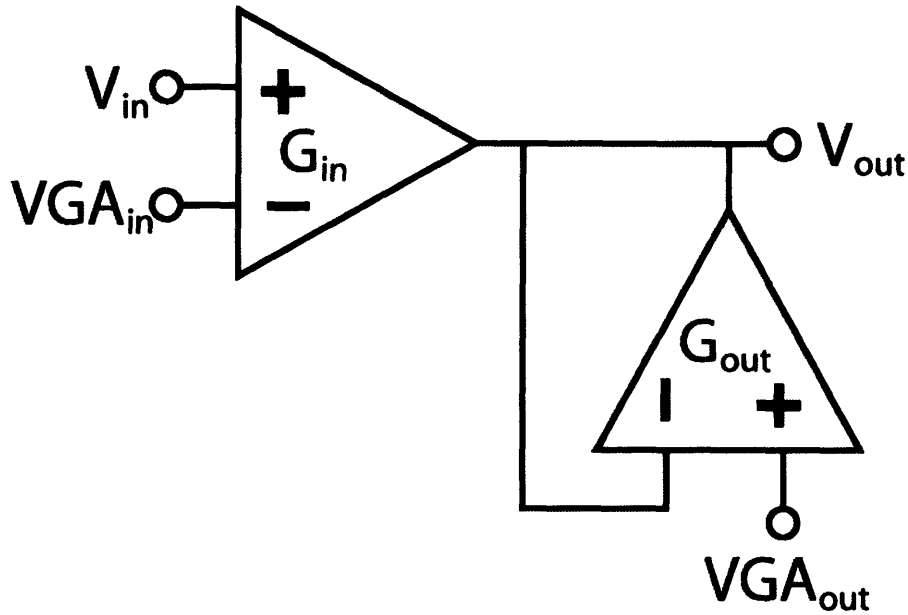
Figure 4.14: Variable gain amplifier circuit. This is a subset of Fig. 12. Here two transconductors are used to provide a continuously variable gain.

The connections to the VGA distinguish between the compression and expansion half channels. To perform compression the system is connected as shown in Fig. 4.15 with the output of the envelope detector controlling G_{out} so that large signals receive less gain.
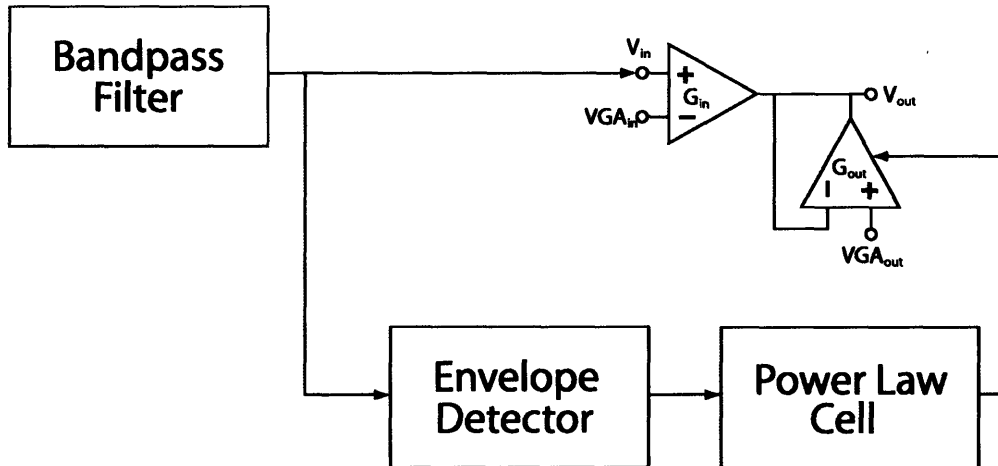
Figure 4.15: Compression half channel block diagram. To perform compression, the output of the power law cell is connected to the output transconductor of the variable gain amplifier. In this configuration a small signal creates a small output transconductance and high gain.

The bandwidth of the VGA is determined by G_{out} and the load capacitance. The load capacitance is the combination of the input capacitance for G_{out} and the input capacitance of a

138

source follower that buffers the output. This total capacitance is approximately 200fF. To ensure that bandwidth exceeds 4kHz

$$\frac{G_{out}}{C} \geq 2\pi 4000 \tag{4.57}$$

$$G_{out} \geq 12000(200fF) = 2.4nS \tag{4.58}$$

$$G_{out} = \frac{I_{out}\kappa}{4\phi_t} \tag{4.59}$$

$$I_{out} \geq \frac{2.4nS(0.1)}{0.7} = 0.35nA \tag{4.60}$$

The output current is at its minimum when the gain is at its maximum. This occurs for the smallest signal. To obtain a compression index of 0.25, the input dynamic range of 40dB must be compressed to 10dB. A signal 20dB below the midline input, must be amplified 15dB, a factor of 5.6, to 5dB below the midline output. So,

$$I_{in} = 0.35nA(5.6) = 2nA \tag{4.61}$$

The power estimate for this stage is

$$Power_{VGA,comp} = 2.8(16 channels)(2 + 0.35) = 106nW \tag{4.62}$$

The primary current of the power law circuit for the compression half channel can now be calculated as well.

$$Power_{powerlaw,comp} = 2.8(16 channels)0.35 = 16nW \tag{4.63}$$

These power consumption equations are for the case when the signal is small. As the input signal grows, the gain in the VGA decreases to -15dB. This gain requires

139

$$I_{out} = I_{in}(5.6) = 11.2nA \qquad (4.64)$$

Here the power consumption is

$$Power_{VGA,comp} = 2.8(16channels)(2+11.2) = 590nW \qquad (4.65)$$

and

$$P_{powerlaw,comp} = 2.8(16channels)(11.2nA) = 502nW \qquad (4.66)$$

The noise from the VGA can be calculated using the same method that was used for the filters. The output current noise for the transconductors is $4qI_{bias}$, so the noise current at the output node for a minimum signal is

$$I^2_{noise} = 4q(I_{in} + I_{out}) = 4q(2.35nA) = q(9.4 \cdot 10^{-9}) \qquad (4.67)$$

The output transconductor converts this to a voltage of:

$$V^2_{noise} = \frac{I^2_{noise}}{G^2_{out}} = \frac{9.4q(150mV)^2}{(0.35nA)^2} = 2.7 \cdot 10^{-10} \frac{V^2_{rms}}{Hz} = 17 \frac{\mu Vrms}{\sqrt{Hz}} \qquad (4.68)$$

Integrated over the channel bandwidth leads to a total noise of 251μVrms. Integrating over the system bandwidth, the noise is 1.1mVrms before it is filtered by the narrow band filter. These are both more noise than calculated for the bandpass filter, but this is the noise present in the compressed signal. It is more useful to consider the noise in the uncompressed signal by using input referred noise.

$$V^2_{noise} = \frac{I^2_{noise}}{G^2_{out}} = \frac{9.4q(150mV)^2}{(2nA)^2} = 8.4 \cdot 10^{-12} \frac{V^2_{rms}}{Hz} = 3 \frac{\mu Vrms}{\sqrt{Hz}} \qquad (4.69)$$

That is 43μVrms in the channel and 183μVrms over the entire bandwidth of the system.
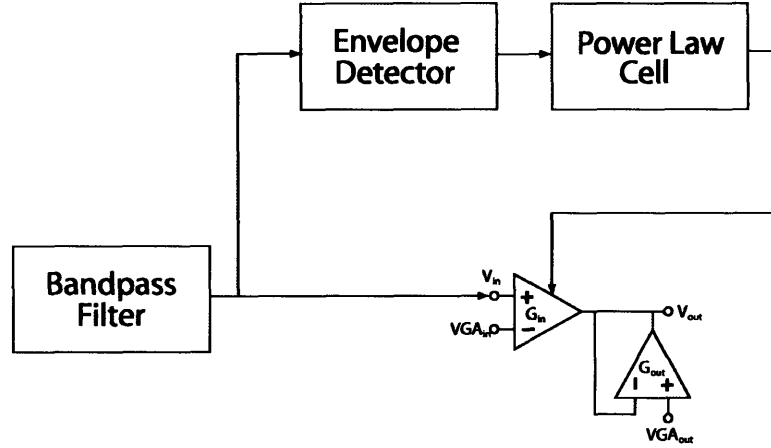
Figure 4.16: Expansion half channel block diagram. The expansion half channel differs from the compression half channel only in the connection of the output of the power law cell to the variable gain amplifier. In the expansion half channel, a small signal creates a small input transconductance in the variable gain amplifier leading to a small gain.

In the expansion half of each channel, the gain is controlled by adjusting the current to the input transconductor. With this topology, the gain increases as the signal increases. The bandwidth is constant because the output transconductance is constant. The output current could be set at 0.35nA as shown above to preserve the bandwidth, but in this design the output current was set to the same 2nA that was used to bias G_{in} in the compression half channel to lower the noise and maintain symmetry between the two halves of the channel. As in the compression half, the smallest signal generates the smallest current and the power consumptions are the same for this half of the channel.

The output current noise is also the same as for the compression channel, but G_{in} and G_{out} have been switched so the noise for the expansion half channel block is

$$V^2_{noise} = \frac{I^2_{noise}}{G^2_{out}} = \frac{9.4 \cdot 10^{-9} q (150 mV)^2}{(2nA)^2} = 8.4 \cdot 10^{-12} \frac{V^2_{rms}}{Hz} = 3 \frac{\mu Vrms}{\sqrt{Hz}} \qquad (4.70)$$

That is 44μVrms in the channel and a total of 184μVrms over the system bandwidth.

Section 3. System Design

Each of the sixteen channels has sixteen adjustable parameters for a total of 256 parameters for the chip. This presents two system level problems. First, the values must be adjustable.

Second, the effects of each parameter must be independently measurable so variations inherent in analog design may be compensated for.

This chip uses two methods for adjusting programming. Inside the channels, the conservative method of resistive biasing shown in Fig. 4.17 is used. Current sources in each channel are biased by voltages which are divided using a resistive ladder. The exponential relationship between voltage and current creates the desired log spacing of bias currents.
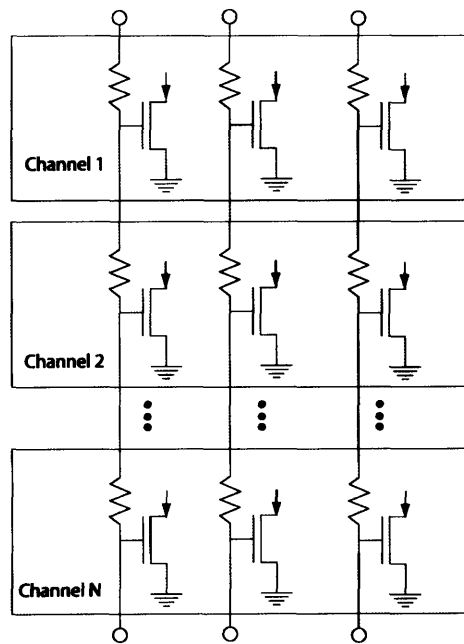

Figure 4.17: Resistive divider biasing. This conservative technique uses voltages to bias transistors used as current sources at each channel. Resistors placed along the line create a voltage divider. The constant voltage division between the channels leads to an exponential division of currents.

At the output of the system, a more advanced programmable DAC system is used to program a weighted sum circuit shown in Fig. 4.18. The transconductor at the output of each channel creates a current proportional to the difference between that channel and the system output. The DAC bits are programmed through a three wire interface that is designed to support DACs for all of the parameters on the chip. Two of the wires are used to shift bits into the global programming shift register shown in Fig. 18. When the register is full of valid data, the third wire is raised high. This enables the channel

142

level logic. Each channel compares the channel address in the global programming shift register with its internally stored channel address. The data fields are then latched in to the DAC with the corresponding DAC address.
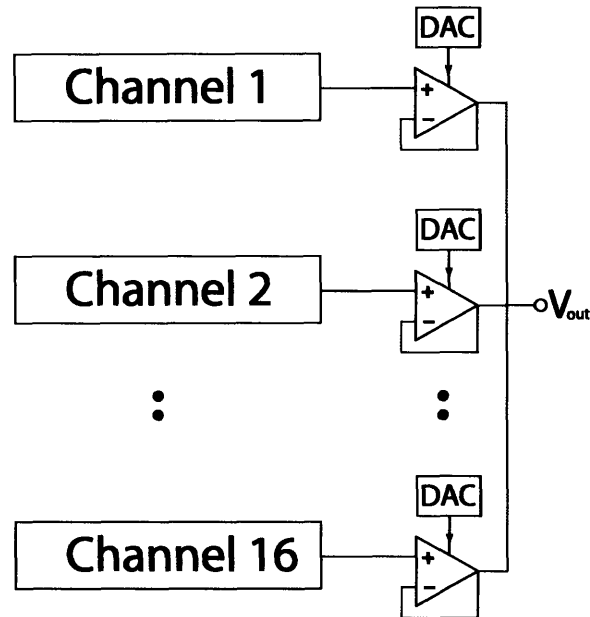

Figure 4.18: Programmable channel summer. The transconductor at the output of each channel creates a current proportional to the difference between the channel output and V_{out}. This creates an average of the channel outputs weighted by the DAC values.
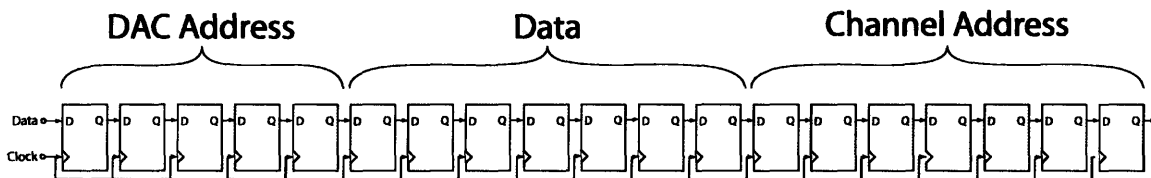

Figure 4.19: Programming shift register. The programmable shift register reduces the number of wires required to program the chip. Data is shifted into the register from a data pin using a clock pin for timing. A third pin causes the chip to store the data in the latch associated with the correct DAC and channel.

This system solves the visibility problem by placing each stage in a constant gain mode. This technique can be demonstrated for the case of measuring the frequency response of the input filter. The variable gain amplifier becomes a constant gain amplifier if it is driven with a constant input current. This can be guaranteed by setting G_{in} of the power law circuit to zero. Then the output filter can be made flat for frequencies of interest by setting its Q to an arbitrarily low value. Then the output of the channel provides an accurate measure of

amplitude changes in the first filter. By biasing the system one block at a time, each block can be characterized.

Section 4. Single Channel Experimental Results

The sixteen channel system depicted in Fig. 4.20 was fabricated through MOSIS on the AMI 0.5μm process. The sixteen channels are stacked from top to bottom. The signal enters all of the channels on the left. On the right, both expansion half channel envelope currents and signal voltages are brought out from each channel. Biasing voltages are provided at the top and bottom.
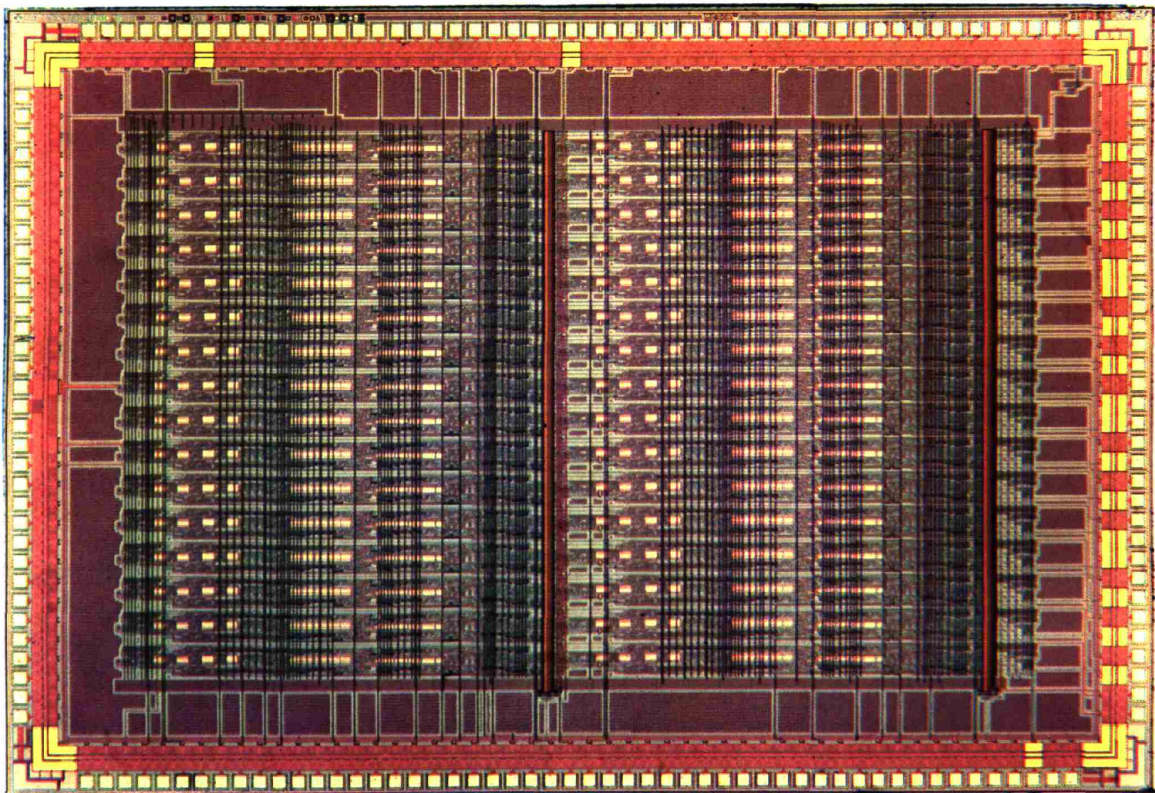

Figure 4.20: Companding system chip die photograph. This chip was built on AMI's 0.5um process through MOSIS.

The channel was tested by first testing the filter block (subsection A). Then the envelope detector was tested (subsection B). Next the compression half channel was tested (subsection C). Finally, a single channel was tested with compression and expansion (subsection D).

A: Filters

The filters were first tested for functionality. The frequency sweep shown in Fig. 4.21 was measured at the output of the input filter using a Stanford Research SR720 lockin amplifier. The sweep demonstrates second degree roll-off on both the high and low side. The filter is well behaved for values of q_f up to 8, but at higher q_f the problem demonstrated in the frequency response in Fig. 4.22 develops. As shown in Fig. 4.9, the filter is formed as the cascade of two first order seconds. The two sections are biased with mirrored currents, creating the possibility for mismatch from both the current mirror mismatch and mismatch in the kappa of the two transconductors. Approximating this mismatch as ten percent, the maximum q_f of the filter for which the center frequencies will fall within the 3dB bandwidth is approximately

$$0.1 \le \frac{1}{q_f} \tag{4.71}$$

$$0.1 q_f \le 1 \tag{4.72}$$
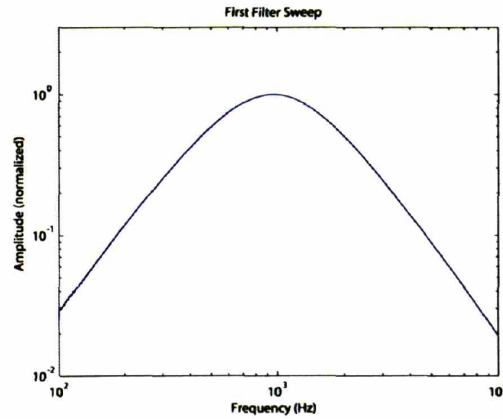
$$q_f \le 10 \tag{4.73}$$

Figure 4.21: Filter frequency response. The bandpass filter is well behaved with second-order roll-off on both sides.
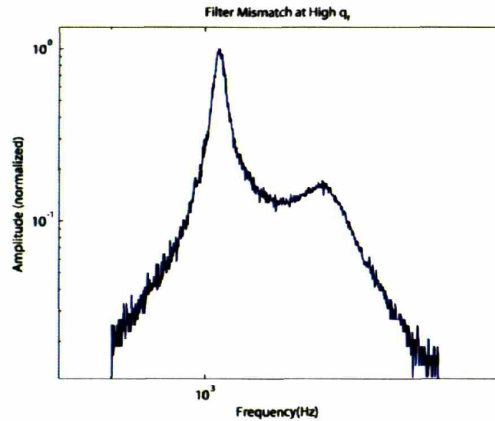
Figure 4.22: Filter frequency response at high values of q_f. Mismatch between the two stages of filtering degrades filter performance when $q_f > 10$.

Noise was measured immediately after the filter using a Stanford Research SR785 spectrum analyzer. Figure 4.23 demonstrates that the noise fit is good in the passband, but the experimental setup has excess noise at higher frequencies. Integrating the excess noise in this plot shows that it nearly doubles the noise in the channel from 200μVrms to 350μVrms. Referring back to Eq. 4.31, only the final buffer contributes noise at these frequencies. But to fit the noise data, the noise from the buffer must be five times as large as calculated. This additional factor is caused by noise in the bias current for the buffer. Unlike differential transconductors, where bias current noise is removed by common mode rejection, the buffer transmits the noise present in its bias current. Removing the factor of five by adding filtering capacitors to the bias current circuitry would reduce this buffer noise from 150μVrms to 30μVrms.
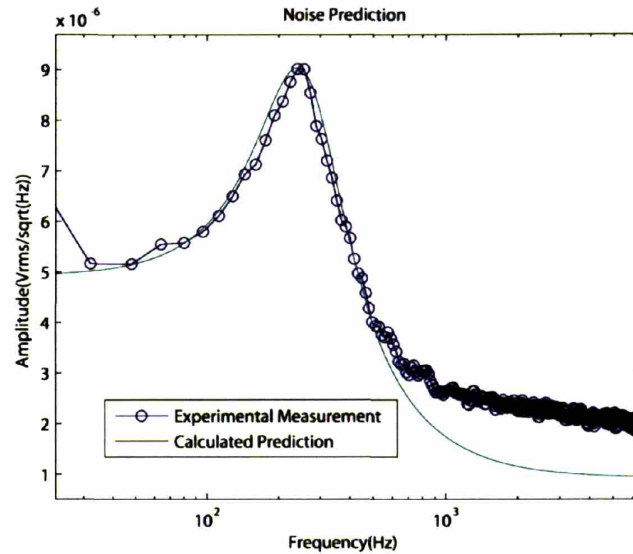
Figure 4.23: Filter noise. The experimental measurement agrees well with the calculated prediction in the passband, but the output buffer introduces excess noise at high frequencies.

At the other end of the amplitude spectrum, distortion was measured using the lockin amplifier to measure the first three harmonics. Figure 4.24 shows that for a q_f of 2, the distortion grows rapidly at 50mVrms.

Figure 4.24: Distortion measurement. The linear range is limited by both second and third harmonics when the fundamental amplitude is approximately 50mVrms.

147

The maximum signal defined by 5% total harmonic distortion decreases with q_f as predicted and as shown in Fig. 4.25. It also shows that noise is relatively invariant with q_f, so the dynamic range falls from 46dB at a q_f of 2 to 17dB at a q_f of 18.

Dynamic Range versus q_f

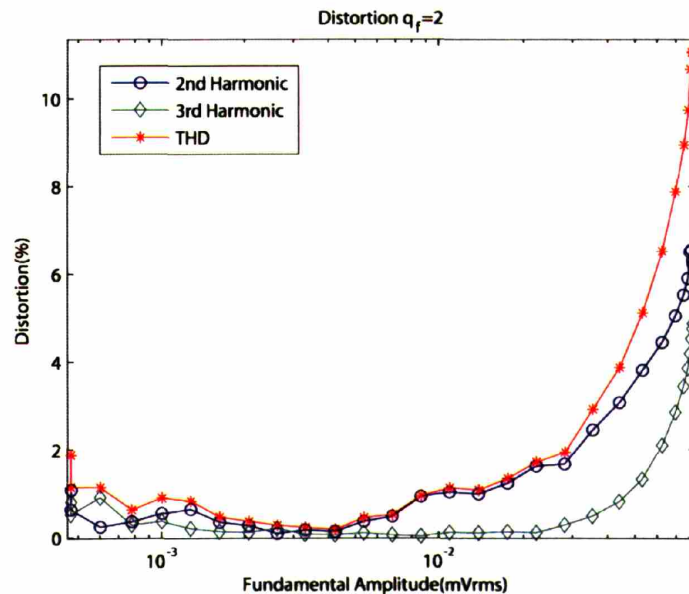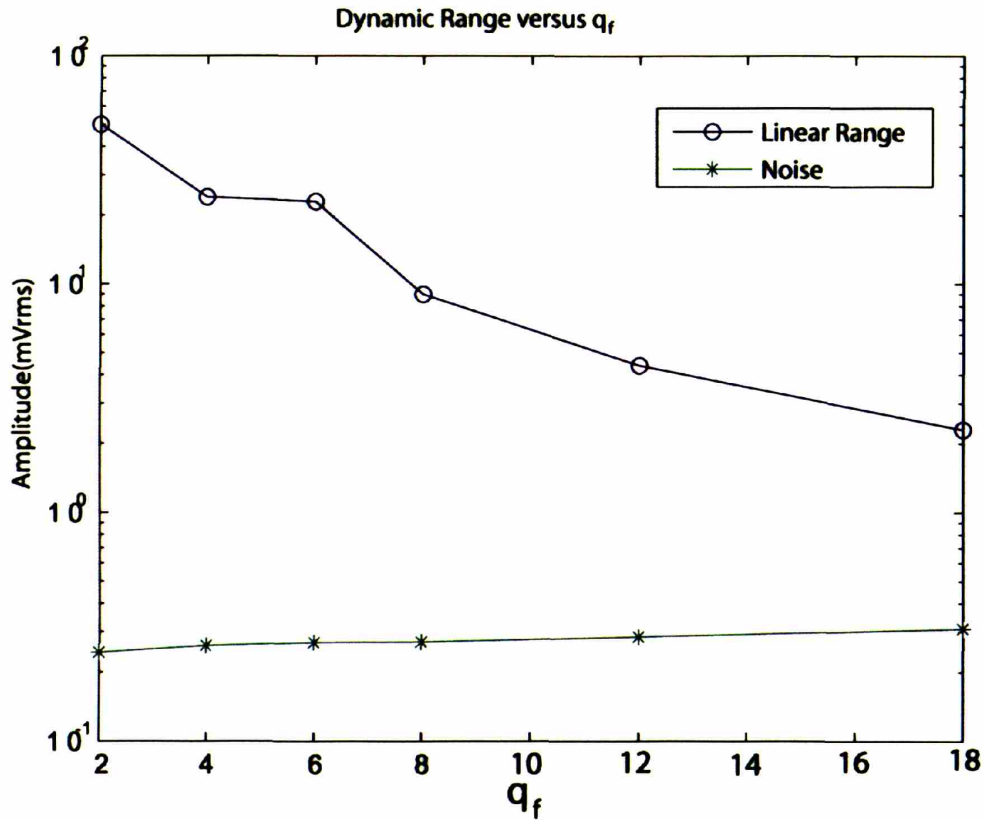Figure 4.25: Dynamic range versus q_f. The dynamic range is the difference between the noise and the linear range. This plot shows that the noise changes very little with q_f while the linear range decreases significantly.

The power of the filter was measured by measuring the increase in system power when all sixteen filters were biased at the voltages that create 1kHz at q_f=2.8 in the observed channel. The measured power was 175nW compared to the designed value of 95nW.

B: Envelope Detector

The output envelope detector was used for characterization because its current was accessible. The current was measured using a Keithly Instruments 6420 Electrometer. Greater than 40dB of linearity is shown in Fig. 4.26.
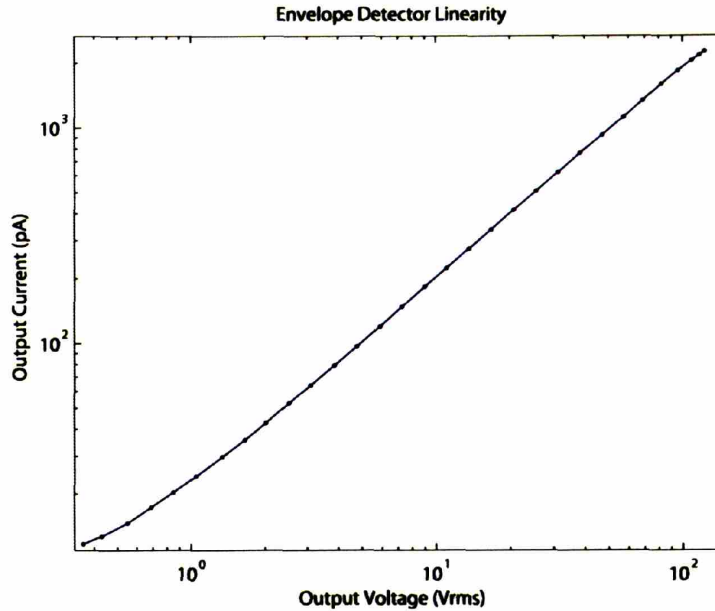
Figure 4.26: Envelope detector linearity. The output current of the envelope detector is linear with the input voltage amplitude over greater than 40dB variation.

As predicted by the analysis, the linearity is a function of bias current and frequency. Figure 4.27 shows the envelope detector output current as a function of input amplitude and bias voltage at 250Hz. The system begins to lose dynamic range between the middle and lower curve, between 1nA and 300pA. At higher frequencies, the deadzone is more significant since the current has less time to charge the parasitic capacitance. Figure 4.28 shows that at 5kHz the deadzone effect is evident even with 3nA of bias current, but that it continues to demonstrate 40dB of range at the 3nA bias. The envelope detector achieved best performance while running at 133nW compared to the designed value of 112nW.
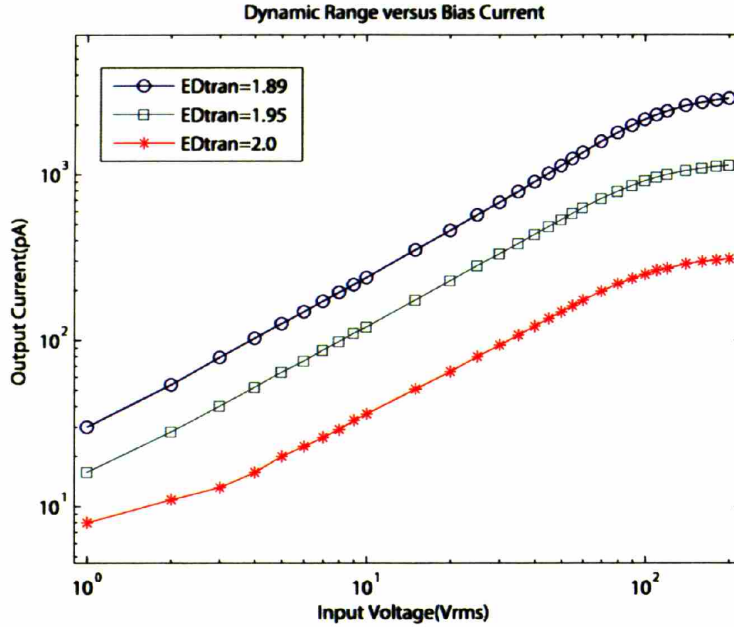
Figure 4.27: Dynamic range at 250Hz. The input-output characteristics of the envelope detector are plotted for three different values of the input transconductance. The effect of the deadzone at small amplitudes can be seen only in the lowest bias condition when ED$_{tran}$=2.0V and the transconductor is biased at approximately 300pA.
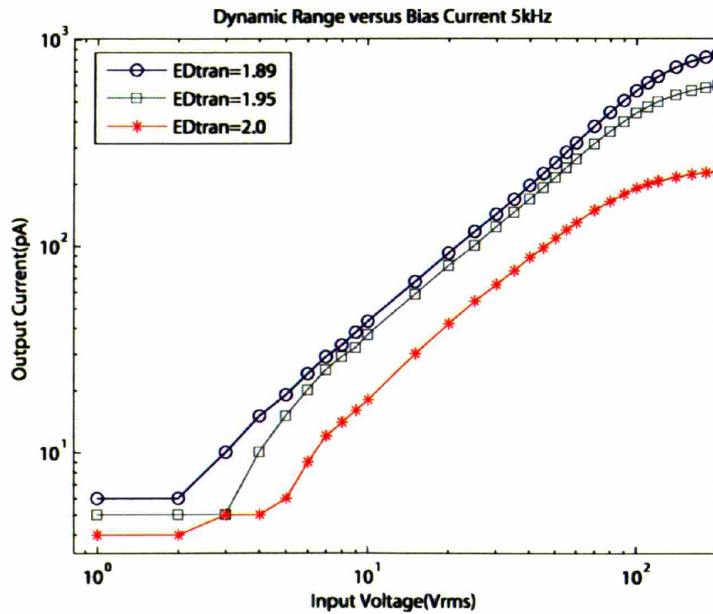
Figure 4.28: Dynamic range at 5kHz. The same experiment as Fig. 26 was performed here with the channel frequency at 5kHz. The higher frequencies makes the deadzone a bigger problem.

The basic functionality of the peak detector is required for all of the envelope detector measurements, but the dynamics of the peak detector can be further measured by using the

150

entire compression half channel. The system is driven with a sine wave modulated by a square wave as shown in Fig. 4.29.
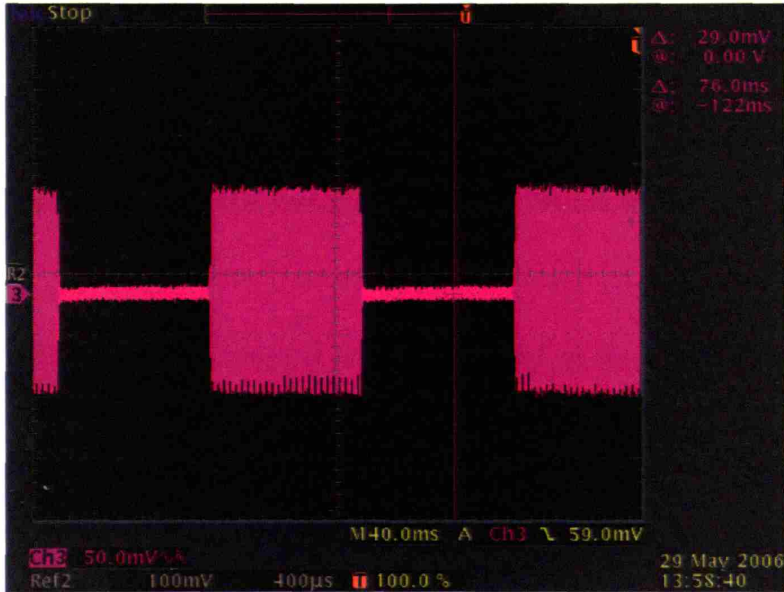
Figure 4.29: Square wave modulated sine wave. This signal is used to test the dynamic response of the peak detector.

The response time of the system, shown in Figs. 4.30, is dominated by the attack and release time constants. The oscilloscope cursors are measuring the fifty percent rise time. This is related to the time constant by

$$0.5 = e^{-\frac{t_{50}}{\tau}} \qquad (4.74)$$

$$t_{50} = -\tau \ln 0.5 \approx 0.69\tau \qquad (4.75)$$

So the time constants shown in the figures are 5, 10, and 20ms.

151

(A)   **(B)**   **(C)**
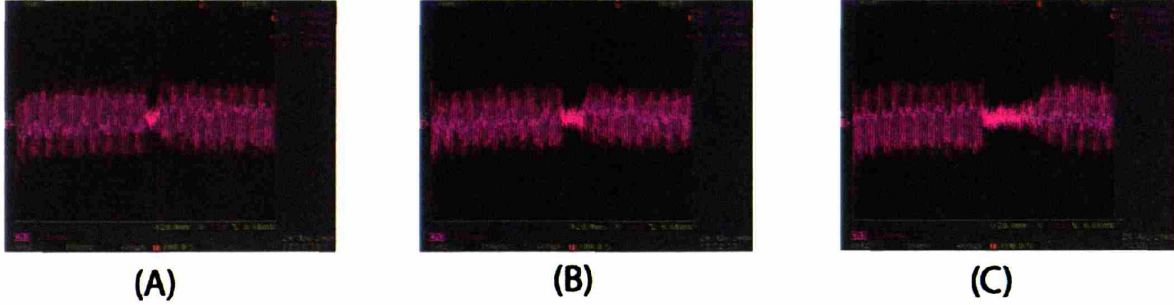
Figure 4.30: Peak detector release times. This figure shows the output of the compression half channel. The system is compressing the input to a constant amplitude. In response to a step down in amplitude, the system increases the gain as the peak detector output decays. This provides a measurement of the peak detector release time constants. The time constants in this figure are: 5ms (A), 10ms (B), and 20ms (C).

The power consumed in the peak detector is 3.3nW, but there is significant excess power associated with creating the reference voltage used in this topology. V_{ref} in Fig. 4.12 must be a low impedance voltage. This was created using the same buffer biased at 100nA as was used in the bionic ear design [8]. This power should be reduced in future designs.

C: Compression Half Channel

On the first channel, the output of the compression half channel can be measured directly. Figure 30 demonstrates the adjustment of the companding index, n_1, from 1 to 0.5. The curves cross at the point where $I_{ref,in}$ is the input amplitude and $I_{ref,out}$.

152

Compression Curves
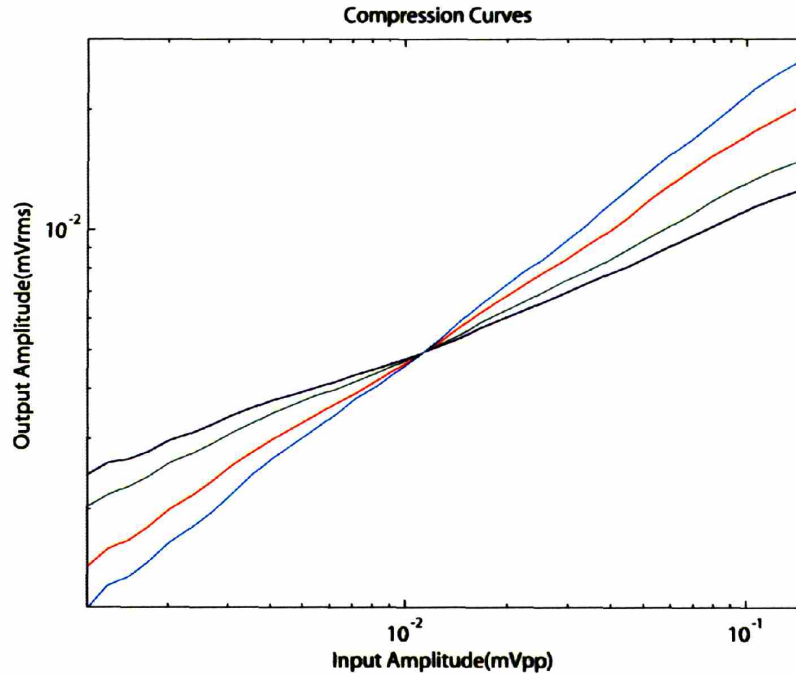
Figure 4.31: Adjustable compression curves. The output amplitude was measured using a lockin amplifier while the input amplitude was swept. The four curves were taken with different values of the companding index, n_1, from 1 to 0.5.

The system was design for this pivot point to be placed at the middle of the dynamic range. If the pivot is placed too low, the power law circuit can move outside of its linear range as analyzed in section and the variable gain amplifier can be biased in a poor bandwidth condition as analyzed in section . The combined effect is demonstrated in Fig. 4.31.
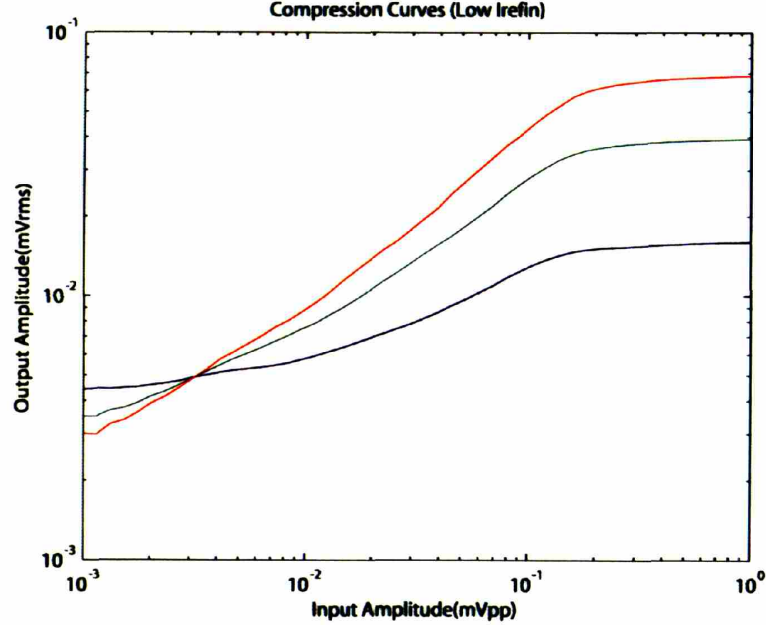
Figure 4.32: Compression curves measured with a low pivot point. Compression curves were measured as in Fig. 30, but with $I_{ref,in}$ set to a low value. This biasing lowers the pivot point of the compression curves and limits the linear range of excursions on the high side. This nonlinearity shows up in the curvature of the curves even at low amplitude.

D: Single Channel

The complete single channel performance was demonstrated by recreating the two-tone suppression figures from the original companding paper[2]. In each of these figures, the system was driven with the sum of two sinusoids

$$v_{in}(t) = a_1 \sin(\omega_1 t) + a_2 \sin(\omega_2 t + \phi_0) \tag{4.76}$$

The power in the output of the channel was measured using the Stanford Research SR785 spectrum analyzer.

For the first experiment, Fig. 4.33, both tones were placed at frequencies within the broad filter while only the first tone was in the passband of the narrow filter. For three different values of suppressed tone amplitude, a_1, the suppressor tone amplitude, a_2, was swept. The dynamic range of the experimental system does not allow the 200dB output range shown in the simulation, but the experiment recreates two important properties of the simulation. First, the suppression level grows linearly in dB with the suppressor tone level. Second, the three curves

154

are linear translations of each other, i.e. the break point amplitudes and suppression slopes are equal.

Figure 4.33: Two-Tone Suppression as a Function of Suppressor Amplitude. (A) is a plot from simulation taken from [2]. (B) is experimental data from this chip. The amplitude of a_1 was kept constant for each curve while the amplitude of a suppressor tone, a_2, was swept. As the suppressor tone increase in amplitude, it suppressed the test tone.

The simulation and experiment in Fig. 4.34 were performed with tones at the same two frequencies as in Fig. 4.33. Here the companding ratio, n_1, is set to three different values. The slopes of the compression curves measured at the middle of the channel vary with the companding ratio.
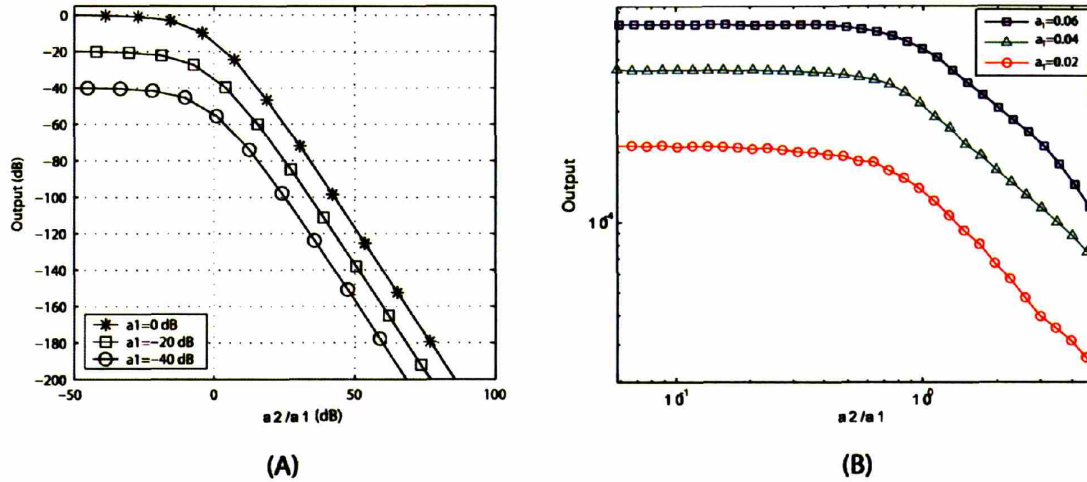
155
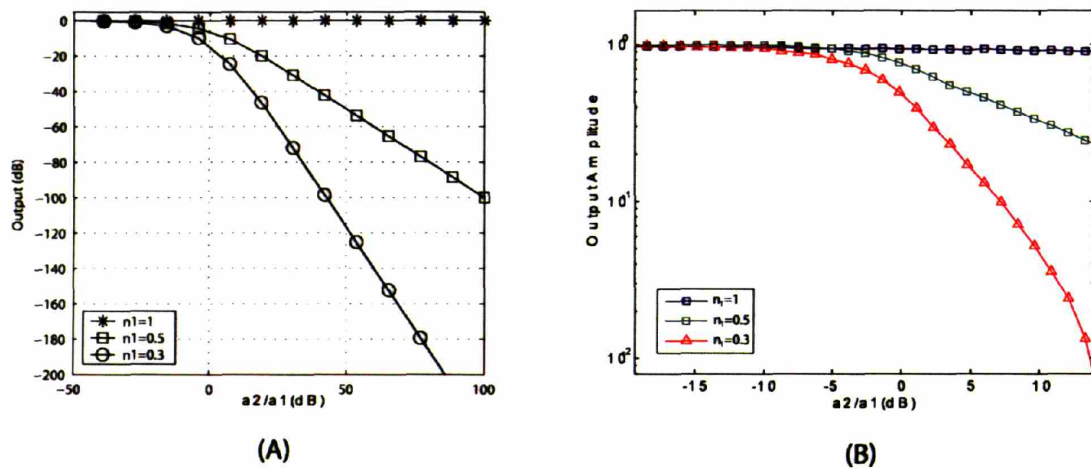
Figure 4.34: Two-Tone Suppression Versus Suppressor Amplitude. (A) is a plot from simulation taken from [2]. (B) is experimental data from this chip. The degree of suppression was determined by the companding index. n_1.

The system can also be characterized by sweeping the suppressor tone frequency rather than the suppressor tone amplitude. Plots such as those in Fig. 4.35 demonstrate the effects of both the broad and narrow filter. In these plots, the amplitude of the suppressor tone is ten times the amplitude of the suppressed tone. At both very low and very high frequencies, the suppressor tone has no effect and the suppressed tone determines the output power. At intermediate frequencies, the suppressor tone passes through the broad filter, but not the narrow filter. At these frequencies, the power at the output is reduced. At the center frequency, the larger suppressor tone is passed through both filters and the output power reaches a maximum. Figure 4.35 demonstrates that the bandwidths stay the same, but the level of suppression at each frequency varies with the compression index.
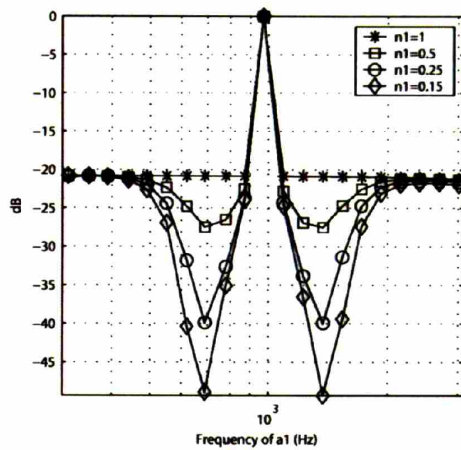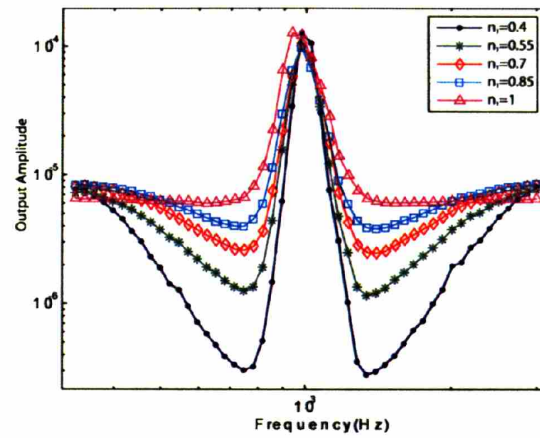
(A)   (B)

Figure 4.35: Two-tone suppression versus frequency. (A) is a simulation from [2]. (B) is experimental data from this chip. The power at the output of the channel was measured while the frequency of the suppressor tone was swept. When the suppressor tone was far from the center frequency the suppressed tone set the amplitude. As the suppressor approached the suppressed tone, suppression reduced the power at the output. Then when the suppressor tone was within the narrow filter bandwidth the powerful suppressor tone was transmitted to the output causing the peak.

In both simulation and experiment in Fig. 4.36, expanding the bandwidth of the broad filter increases the range of frequencies which lead to suppression. The center response is constant while the suppression curves differentiate at different levels.
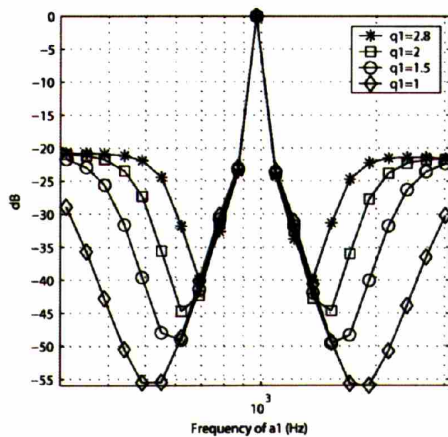

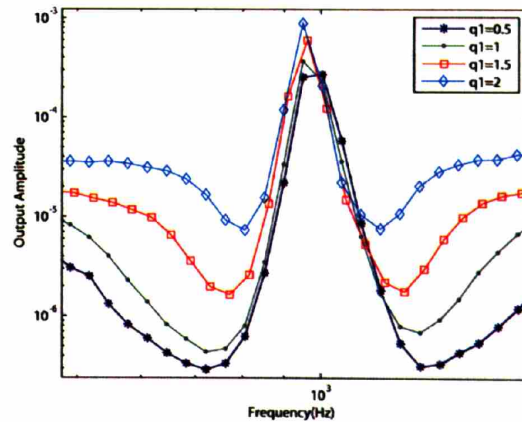(A)   (B)

Figure 4.36: Two-tone suppression versus q_1. (A) is a simulation from [2]. (B) is experimental data from this chip.

157

The center response was adjusted for the simulation and experiment shown in Fig. 4.37. Here the surround curves are constant while the center curves vary in bandwidth. In the experiment, the peaks are attenuated at higher q_f values.

(A)   (B)
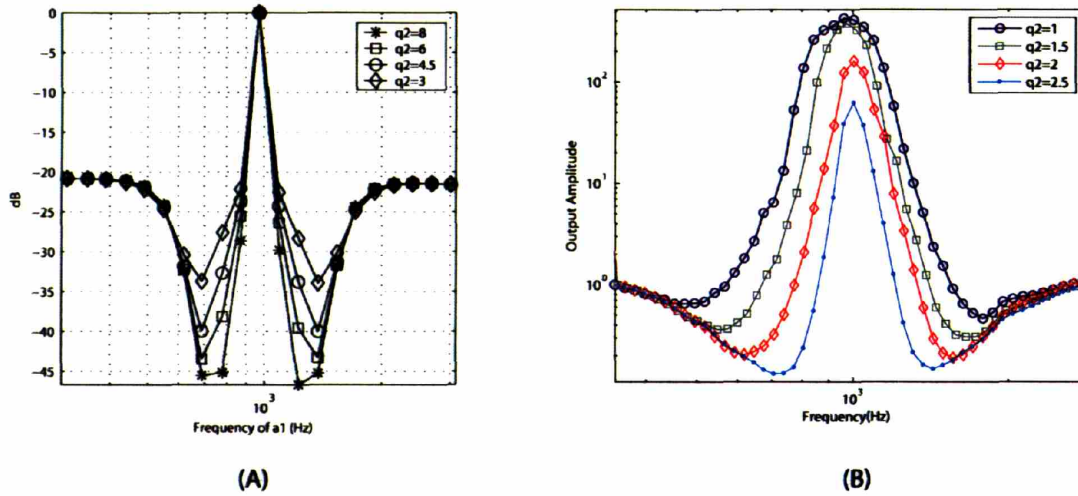
Figure 4.37: Two-tone suppression versus q_2. (A) is a simulation from [2]. (B) is data from this chip.

The half channel has a low power consumption of less than $2\mu W$ as shown in Table 1. The total power for the chip is approximately $60\mu W$. Each of the currents was biased slightly higher than the designed value. In future designs, the power should be cut in half by minimizing the power consumed in buffers. Each half channel has two buffers, the peak detector reference buffer and the signal output buffer, which were conservatively biased.

Component	Designed Power	Measured Power
Filter (1kHz, q_f=2.8)	147nW	175nW
Filter (1kHz, q_f =4.5)	210nW	218nW
Peak Detector	0.8nW	3.3nW
Envelope Detector	112nW	133nW
VGA	6.6nW	31nW
Buffers	1120nW	1320nW
Half Channel Total	1.6μW	1.9μW

Table 1: Power Consumption

Low power consumption is key to making companding applicable to hearing instruments. A standard hearing aid type 312 battery provides 175mWh of power in a form factor used in both in the ear and in the canal hearing aids [2]. If this battery was used to power a companding

158

speech processor, it would last 6 months with 16 hours of use per day. In practice, hearing aid batteries are replaced once a week. In that time, the companding speech processor uses 3.8% of the battery's capacity.

Section 5. System Performance

The system chip is composed of sixteen identical channels corresponding to the sixteen channels in a typical cochlear implant. Combining multiple channels allows the system to create a companded spectrum. Noise suppression is shown in this companded spectrum in Fig. 4.38. The system was first driven with white noise to generate the reference waveform. Then it was driven by a single tone at 1kHz added to the white noise. The tone suppresses the noise at adjacent frequencies.
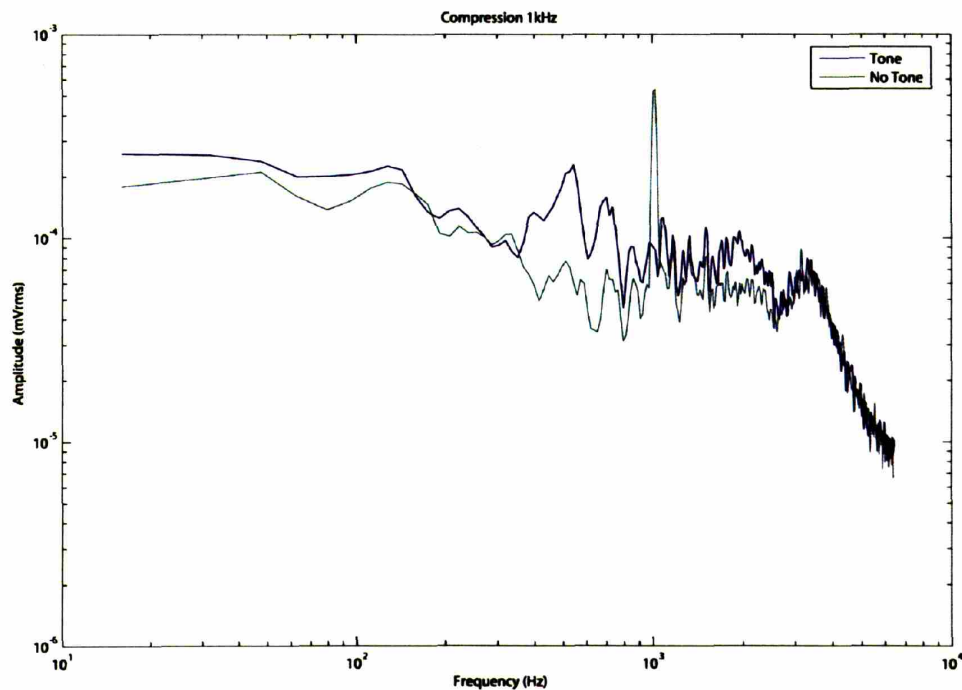

Figure 4.38: Noise Suppression From 1kHz Tone.

The performance of this system is limited by two primary constraints not found in the channel. Section I discusses the problems caused by adding up the noise from multiple channels. Section II discusses the problems with biasing caused by mismatch between the channels.

159

A: Noise

The signal at the output of each channel is confined to a narrow frequency band by the narrow band filter in the expansion half channel. This filter also limits the noise of the earlier stages of processing, but the noise from two broadband stages are not included in this filtering. The first stage is the buffer placed at the output of the bandpass filter. This is the buffer responsible for the excess high frequency noise in Fig. 21 and accounts for up to 150μVrms of out of band noise with current biasing or 30μVrms predicted from filtered biasing. The second broadband noise source is the VGA. Because of the low currents at which the VGA is run, it contributes more noise. As shown in section 2.D., the VGA has 184μVrms of out-of-band noise. This out-of-band noise adds between the channels, scaling by $\sqrt{channels}$, to 736μVrms for 16 channels.

The solution to this out of bound noise is to ensure that all of the noise is filtered. This can be accomplished in two ways. A new filter may be added following the buffer and VGA or the buffer and VGA may be included in the existing narrow band filter.

The simplest solution is to add an additional stage of filtering following the VGA. In addition to the added area and power requirements of adding an additional filter, this filter will be difficult to implement because it must operate on the full dynamic range of the signal following the expander while filtering with a small bandwidth.
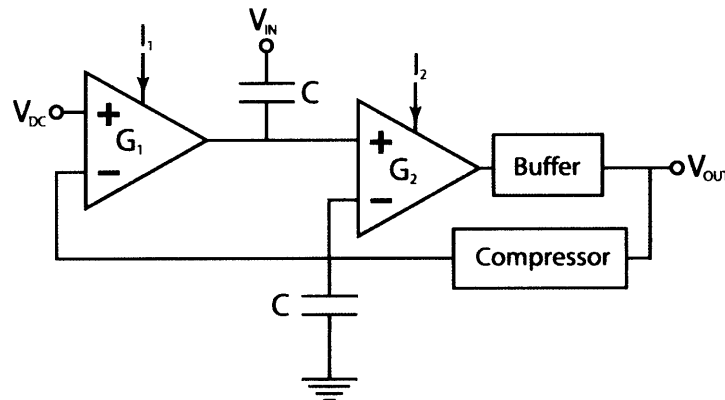
Figure 4.39: Revised Expansion Topology

The second solution simplifies the filtering operation and minimizes components, but is complicated by the required temporal response. The revised expansion is shown in Fig. 4.39. The variable gain amplifier is place in the feedback path of the filter. In this configuration, the VGA attenuates the signal to create a net amplification by the block. This system works for slow frequency changes in the envelope, but to completely recreate the modulation step responses shown in Figs. 4.30 would require an envelope detector capable of instantaneous envelope detection and a slow adaptation. So, further simulations with this topology are required to determine its suitability for this application.

B: Mismatch

The resistive divider system described in Section 3 allows the biasing of the chip to desired values, but it cannot correct for mismatch between the channels. The extent of the mismatch problem is shown in Fig. 4.40. The center frequency and Q are plotted for each channel for five different chips.

Figure 4.40: Filter Mismatch

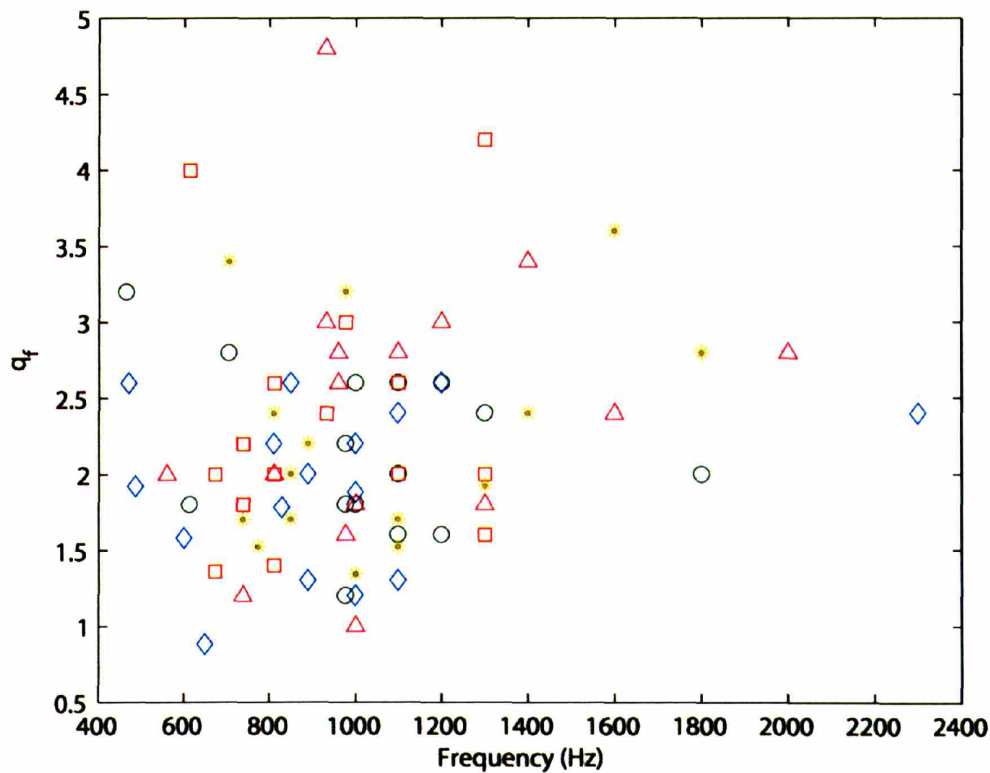For both center frequency and Q the standard deviation is approximately 30% of the mean. The mismatch is even larger for the companding index where it reached 38% of the mean. The mismatch is large in both cases because of the multiple current mirrors involved. For the filter parameters the control currents are first mirrored across the chip by the biasing circuit. Then, the currents are mirrored for use in the translinear biasing. The translinear loops introduce error for internal mismatch as well as mismatch in the q scaling current. Then the output of the translinear circuit is mirrored to the filters. Finally, the conversion from currents to transconductances can be mismatched. There are a total of six contributions to mismatch.

Expanding the DAC based programming system presented in section 3 can compensate for this mismatch. The key remaining parameter is the number of bits to use in the DACs. Data must be taken to measure the required matching for different parameters in the system.

Section 6. Conclusion

The companding algorithm has the potential to improve hearing in noise for hearing aid and cochlear implant users. Turicchia and others have implemented the algorithm in Matlab and begun to explore its performance, but to integrate it into hearing aids and cochlear implants it must be implement in a micropower technology. I have built a custom analog integrated circuit that performs the companding operation in less than $60\mu W$. At that power level, the companding speech processor could run for 6 months from a standard hearing aid battery.

I have shown that my micropower analog companding processor exhibits two-tone suppression and performs companding. I have also found some of the areas in which the analog implementation differs from the computer implementation and quantified these effects. Now, algorithm can be refined to include the effects of noise and mismatch.

References

[1] K. N. Stevens, Acoustic Phonetics The MIT Press, Cambridge, Ma 1998

[2] H. Dillon, Hearing Aids Boomerang Press Turramurra, Australia, 2001 pp.2-6

[3] L. Turicchia and R. Sarpeshkar, "A Bio-Inspired Companding Strategy For Spectral Enhancement," IEEE Trans. On Biomed. Eng., Vol. 13, No. 2 March 2005 pp. 243-253

[4] A. Bhattacharya and F.-G. Zeng, "Companding to Improve Cochlear Implants' Speech Processing in Noise," abstr. 2005 Conference on Implantable Auditory Prostetheses, Asilomar, California 2005

[5] R. Sarpeshkar, R. F. Lyon, and C. A. Mead, "A Low-Power Wide-Linear-Range Transconductance Amplifier," Analog Integrated Circuits Signal Process., vol. 13, no. 1/2, May/June 1997.

[6] G. Groenewold, "Optimal Dynamic Range Integrators," IEEE Trans. Circuits Syst. vol 39 iss.8 Aug. 1992 pp. 617-627

[7] J. O. Pickles An Introduction to the Physiology of Hearing 2nd Edition Academic Press Boston, Ma 1988 pp. 277-283

[8] R. Sarpeshkar, C. Salthouse, J-J Sit, M. W. Baker, T.K-T Lu, L. Turicchia, and S. Balaster, "An Ultra-Low-Power Programmable Analog Bionic Ear Processor," IEEE Trans. Biomedical Eng. vol. 52 iss. 4 Apr. 2005 pp. 711-727

[9] C. Mead Analog VLSI and Neural Systems Addison-Wesley Publishing Company, Reading, Mass. 1989.

[10] R. Sarpeshkar, T. Delbruck, and C. A. Mead, "White Noise in MOS Transistors and Resistors," IEEE Circ. And Devices Mag. vol. 9, iss. 6, Nov. 1993 pp.23-29

[11] S. Zhak, M. Baker, and R. Sarpeshkar, "A Low-Power Wide-Dynamic-Range Envelope Detector," Research Lab. Electron Mass. Inst. Technol. Cambridge, Ma, Memo 664, 2003.

[12]S. Zhak, M. Baker, and R. Sarpeshkar, "A Low-Power Wide-Dynamic-Range Envelop Detector," IEEE J. Solid-State Circuits, vol. 38, no. 10, Oct. 2003. pp. 1750-1753

Chapter 5 : Conclusion

I have developed a set of techniques for designing micropower analog filters with both high and low values of Q. First, I demonstrated the capacitive attenuation technique for shifting the dynamic range of gm-C filters to higher linear ranges. Then I demonstrated that for high-Q filters, the dynamic range becomes limited and jump resonance develops. I explored the phenomenon of jump resonance and developed the AQC circuit solution to avoid it. Finally, I implemented the companding algorithm, an architecture that avoids jump resonance while using intentional nonlinearity to process audio signals.

Power efficient filters have been built by other researchers using gm-c topologies. Groenewold then developed a formula for the dynamic range of gm-c filters as a function of filter parameters which can also be used to calculate the power consumption for a given frequency and dynamic range. My work on the capacitive-attenuation filter expanded on this work by demonstrating that attenuators can be used to increase the linear range, while also increasing the noise to keep a constant dynamic range, with minimal increase in power consumption. The technique was demonstrated on a simple second-order filter. This work could be expanded by including alternative filter topologies such as ladder topologies which have demonstrated better device matching insensitivity or by improving the very low frequency behavior of the dc-path transconductor. Currently, the transconductor must be biased at a sufficiently high level to minimize rectification at the common source node.

I worked as part of a larger team to build a micro-power analog speech processor for cochlear implants. I used my capacitive-attenuation filters to separate the input audio signal into sixteen different frequency bands. Our analog speech processor operates on less than 5% of the power of traditional designs.

In order to increase the number of channels in the speech processor, I decreased the bandwidth of each filter, increasing the Q. This decreased the linear range of the filters. In these filters, I observed a number of nonlinear behaviors such as gain compression and harmonic distortion. I also saw jump resonance, a behavior I was not familiar with.

I studied jump resonance using two sets of theoretical methods. First, I used describing function methods to model the non-linearity as an amplitude dependent gain for sinusoidal inputs. With this model, I used the traditional Fukuma-Matsubara method to predict the conditions when can cause jump resonance in my system and my own graphical method to predict the resulting waveforms. I have also developed an intuitive model for the shape of the transfer function based upon positive and negative feedback.

Then, I used a state-space model to explore the behavior of non-sinusoidal waveforms. This analysis demonstrates how the state of the system in the hysteresis region is stored in the voltages on the capacitors. The analysis also predicted a subharmonic mode which I verified in both a circuit simulation and in my experimental setup. This work could be expanded to consider the effect of Q on the regions of attraction, the step response of the system, and the effect of time varying nonlinearities.

Finally, I used my filters in an architecture that avoids jump resonance by using a series of two filters and two discrete nonlinearities with the need for an automatic Q control. These filters and nonlinearities are combined to implement the companding speech processing algorithm as proposed by Turicchia and Sarpeshkar. Preliminary results suggest that the algorithm can improve speech recognition performance for cochlear implant patients in moderate noise environments.

I have demonstrated the key companding speech processing behaviors in my analog processor. A large amplitude tone can cause suppression of an adjacent quiet tone. The level of this suppression and the frequency dependence are fully adjustable.

Future work on the analog companding speech processor should focus on noise and mismatch, two difficulties which are unique to the analog implementation. The noise in the channel can be further reduced by adding filtering to the current bias within the channel. A bigger factor in the noise is caused by out of band noise. When the sixteen channels are added together, the out of band noise is multiplied by 4. An additional filter following the expansion stage would improve this performance. Mismatch can be compensated for by the inclusion of independent digital programming of each of the channel parameters.