Massachusetts Institute of Technology

Department of Electrical Engineering and Computer Science

Submitted in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

Modeling and Design of an Active Silicon Cochlea

by Serhii M. Zhak

Author ___

Date of Submission: September, 2008

Certified by ___

Professor Rahul Sarpeshkar, Thesis Advisor

Accepted by ___

Professor Terry P. Orlando, Chair, Department Committee on Graduate Students

Abstract

Silicon cochleas are inspired by the biological cochlea and perform efficient spectrum analysis: They realize a bank of constant-Q N^{th}-order filters with O(N) efficiency rather than $O(N^2)$ efficiency due to their use of an exponentially tapered filter cascade. They are useful in speech-recognition front ends, cochlear implants, and hearing aids, especially as architectures for improving spectral analysis in noisy environments and for performing low-power spectrum analysis. In this thesis I describe four contributions towards improving the state-of-the-art in silicon-cochlea design, two of which involve theoretical modeling, and two of which involve integrated-circuit design.

On the theoretical side, I first show that a simple rational approximation to distributed partition impedances in the biological cochlea captures its essential features and enables an efficient artificial implementation achieving maximum gain in a minimum number of stages while still maintaining stability. In particular, I show that the terminating impedance of the cochlea is crucial for its stability and discuss various analytic methods for termination. Second, I derive a novel composite artificial cochlear architecture composed of a cascade of all-pass second-order filters from a first-principles analysis of the biological cochlear transmission line. The novel all-pass architecture reduces phase lag and group delay in the silicon cochlea, a problem in prior designs, sharpens its high-frequency rolloff slopes, increases its frequency selectivity, and improves its nonlinear compression characteristics.

On the circuit side, I first present a novel current-mode log-domain topology that simultaneously increases signal-to-noise ratio (SNR) and dynamic range while lowering power consumption in resonant filters with high quality factor Q. The novel topology is validated in a second-order low-pass resonant filter, which is employed in the silicon cochlea, demonstrating a reduction in power consumption and increase in SNR by a factor of Q. When bias currents in the filter are adjusted as the signal level varies, this technique enables an improvement in maximum SNR by a factor of Q and an increase in maximum non-distorted signal power and dynamic range by a factor of Q^4. Measurements from a chip in a 0.18-μm 1.1-V CMOS technology achieve a quiescent power consumption of 580-nW at a 15-kHz center frequency with a maximum SNR of 41.3dB and dynamic range of 76dB for a Q=4. Finally, I describe a current-mode 33-stage 0.18-μm silicon cochlea that achieves 79dB of dynamic range with 41-μW power consumption on a 1-V power supply over a usable 3.5kHz-14kHz frequency range. These numbers represent an 18dB improvement in dynamic range and a 12.5x reduction in power consumption over prior state-of-the-art silicon cochleas.

Thesis Supervisor: Rahul Sarpeshkar

Title: Professor

Acknowledgements

I would like to thank my adviser Professor Rahul Sarpeshkar for providing inspiration, ideas, advice not only in engineering, but in life, and even help on the bench in the lab till 2:30am – truly indicative of his caring. I am tremendously grateful to my thesis reviewers Professor James Roberge and Professor Christopher Shera. Professor James Roberge provided very helpful feedback on the design of the circuits, and Professor Christopher Shera gave many helpful suggestions and ideas about cochlear modeling. And I am tremendously grateful to Professor George Zweig, whom I also had an honor to meet. It is his fundamental paper on cochlear modeling that became the basis for my own theoretical work.

I was constantly surprised by the collegiality of our research group. I am especially grateful to Soumyajit Mandal whose help was absolutely invaluable in everything from my research work to his help, discussions and helpful suggestions on how to make my publications better and to drawing figures for my presentations even! I can not express in words how grateful I am for his help! Dr. Lorenzo Turicchia helped to file my patents, and his suggestions and critical comments proved indispensable for my cochlear modeling. His insights about Italians and life in Italy made my life more joyful, too! Dr. Michael Baker practically introduced me to this research group and was my mentor and friend. I am grateful to my officemate Scott Arfin for all the help with software and hardware alike he provided. Discussions with him also influenced the direction of my research. I will always be eternally grateful to other members of the research group and friends such as Dr. Tim K. Lu, , Dr. Keng-Hoong Wee, Dr. Ji-Jon Sit, Dr. Chris Salthouse, Dr. Heemin Young, Dr. Maziar Tavakoli, Dr. Micah O'Halloran, Scott Arfin, Daniel Kumar, Woradon “Pok” Wattanapanitch, Ben Rappoport, Alex Meevay, Dr. Diana Young. You guys were very understanding, appreciative and helpful in good times and bad, and without you I will definitely not be able to accomplish my graduate study.

The Advanced Bionics Corporation generously provided funding for this project, and my special thanks to Dr. Van Harrison, Dr. Mike Faltys, Salomo Murtonen, Lockshmi Mishra.

My deepest thanks are to my parents, my dad Mykhailo A. Zhak and my mother Galyna G. Zhak. Their love, support, advice, interest in everything in my life from my studies, to my friends, and later even in the deep details of engineering, their pride in me and in my work have set high standard for me very early on in my life. They made me who I am, and I owe them both everything that I have achieved and will ever achieve. They have been always there for me, true friends on every intersection of my life, praising me for successes, but not afraid to frankly point out what needed improvement and suggest how to go about it wisely. I will always love them and will always remember what my roots are!

Finally, thanks to my wife Yuliya for her love, unwavering support and eternal patience through all this time. This is *our* Ph.D.

Table of Contents

1. Introduction

Bionic ears, or Cochlear Implants, have been implanted in more than 20,000 people (Spelman 1999). They mimic the functionality of the ear by stimulating neurons in the cochlea in response to sound. Various algorithms have been employed in bionic ears. The sound is captured by a microphone, divided into frequency bands, then the power in those frequency bands is measured, and finally the neurons are stimulated (Loizou 1998, Ay 1997). In commercially available cochlear implants, a constant-Q wavelet-like bank of bandpass filters is used to decompose the sound signal into frequency bands. But a distributed system of traveling-wave amplifiers is vastly more efficient than a bank of bandpass filters at performing low-power, wide-dynamic-range frequency analysis (Sarpeshkar 2000). One implementation models the system of traveling-wave amplifiers as a cascade of second-order filters with exponentially decreasing corner frequencies (Sarpeshkar 1998, van Schaik 2001). However, a filter cascade is prone to excessive parameter variation sensitivity, noise accumulation and amplification, and also to an accumulation of excessive group delay that complicates the spectrum analysis. This architecture also requires additional filter sections spanning at least one octave to build up the collective amplification. Since the filters in the idle sections operate at the higher end of the frequency range of interest, the system takes a heavy hit in power consumption. Another approach to building an electronic cochlea is by the implementation of transmission-line models. Various kinds of such models were proposed (Zweig 1991, Hubbard 1993, Mammano 1993, Hubbard 2000). The goal of my research is to develop the theoretical aspects of some of the proposed cochlear models with circuit implementations in mind, and to build a low-power wide-dynamic-range active cochlear chip for use in speech processors.

The rest of this Introduction is organized as follows. In section 1.1 we review some previous research on both theory and electronic implementation of the bionic ear. We discuss (1) the standard filter bank architecture of the speech processor used in cochlear implants, (2) a cascade-of-filters architecture emulating active-cochlear operation, (3) passive and active cochlear transmission line modeling and (4) implementation issues. In section 1.2 we present and compare two analog design paradigms, namely voltage-mode, and log-domain or current-mode methods. In section 1.3 we outline the organization of this thesis.

1.1. Background

1.1.1. Mammalian Cochlea

Figure 1 (A) shows the anatomy of the human auditory periphery.

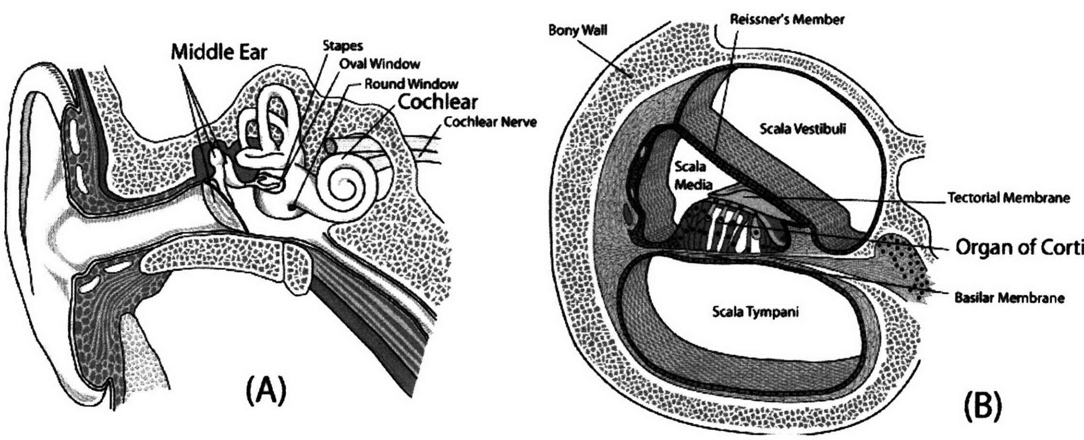


Figure 1: (A) Anatomy of the human auditory periphery; (B) Cross-section through the cochlea. Adapted from (Kessel and Kardon 1979).

Sound waves travel down the canal and vibrate the eardrum of the middle ear. The middle ear serves as an impedance transformer from the low-pressure high-velocity air to the high-pressure low-volume-velocity fluid-filled cochlea. Vibrations of the eardrum couple into the stapes via that transformer. The footplate of the stapes presses on the oval window of the cochlea. The fluid-filled cochlea is partitioned into three compartments, the scala vestibuli, the scala media, and the scala tympani (Geisler 1998) as shown in Figure 1 (B). The oval window displaces fluid in the cochlea and generates a traveling wave of fluid pressure down the length of the cochlea (Dallos 2002). This fluid pressure wave causes displacement of the basilar membrane together with the organ of Corti, which compose a boundary of the cochlear partition (Geisler 1998). The organ of Corti is shown in Figure 2.

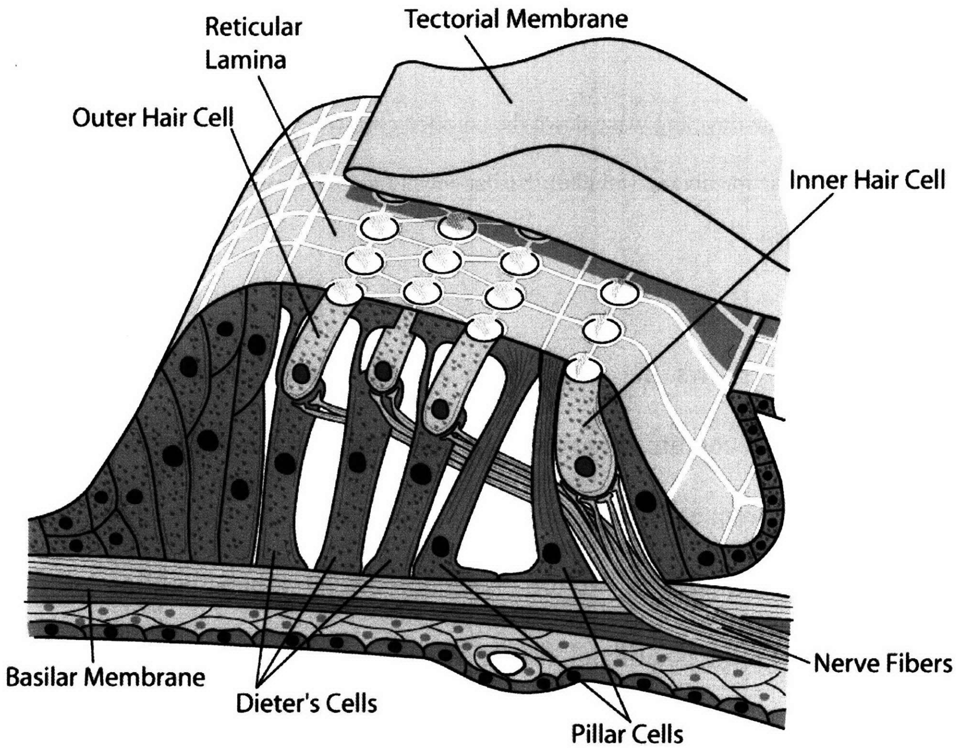


Figure 2: The organ of Corti with the tectorial membrane partially cut away. Adapted from (Kessel and Kardon 1979).

The basilar membrane varies from being light and stiff at the basal end, the end near the stapes, to being heavy and flexible at the apical end. The properties of the tectorial membrane, reticular lamina and outer hair cells within the organ of Corti also vary with the position along the cochlea; the so-called scaling of the organ of Corti's mechanical impedance. As the wave moves from the base to the apex, it resonates with the impedance of the basilar membrane and the organ of Corti peaks at a location that has an associated "best frequency" which matches the frequency of the incoming wave (Dallos 2002). Thus, the cochlea performs a frequency-to-place transformation on the incoming signal.

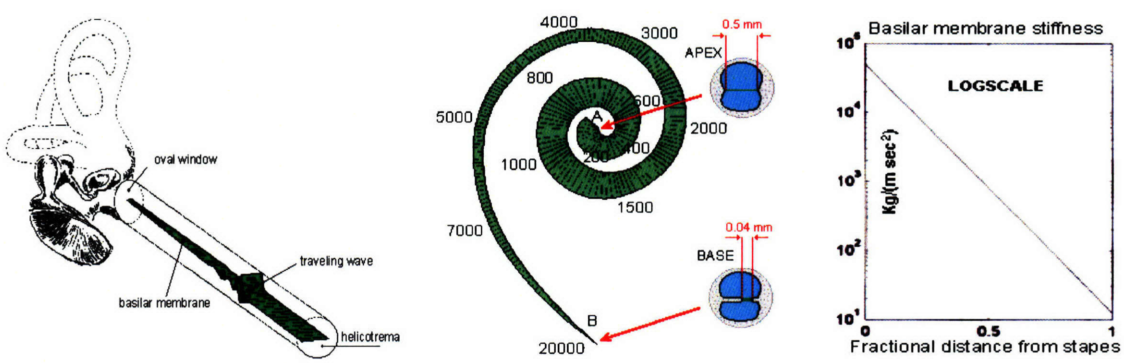


Figure 3: Propagation of the traveling wave down the unrolled cochlear structure (left); approximate frequency map (in Hz) on the basilar membrane (middle); basilar membrane stiffness as a function of the normalized distance from stapes (right).

Figure 3 (left) shows the propagation of the traveling wave down the unrolled cochlear structure by the combined movement of the fluid and the basilar membrane with the organ of Corti. The frequency-to-place analysis performed by such a structure on the incoming signal is illustrated by Figure 3 (middle). The typical scaling of the basilar membrane stiffness with the position along the cochlea is shown in Figure 3 (right).

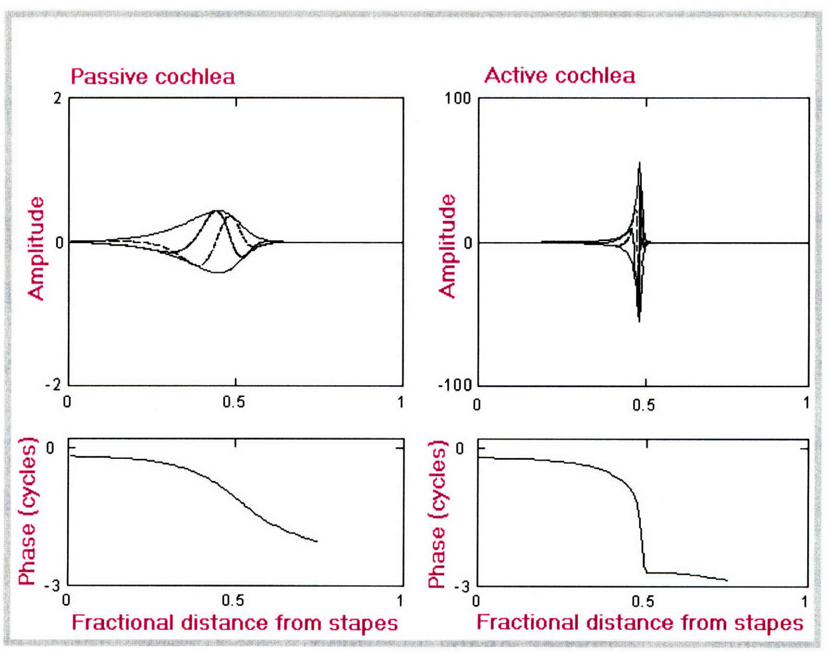
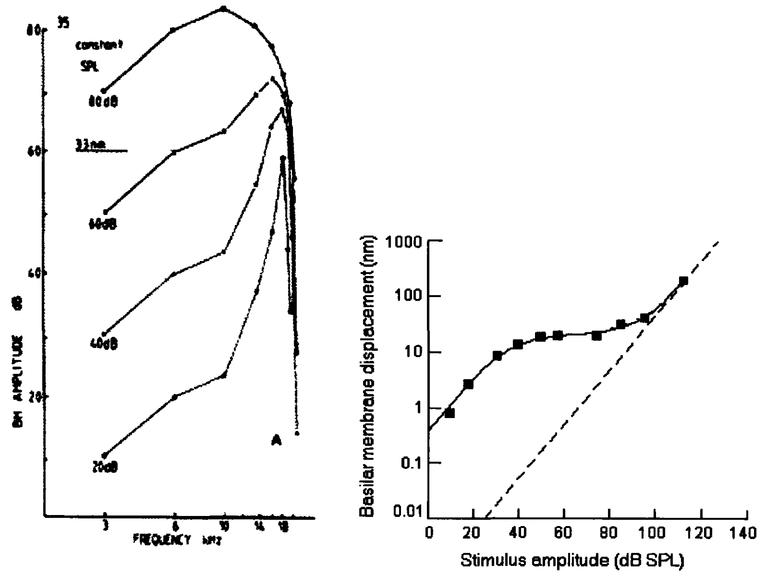


Figure 4: The traveling wave propagation, wave envelopes, and the phase responses for the passive (saturated) cochlea – on the left; and active (alive and unsaturated) cochlea – on the right.



Johnstone et al. (1986)

Figure 5: The non-linear characteristics of the healthy cochlea: The output signal magnitude in dB versus frequency at a fixed position on the basilar membrane at various input sound levels (left); the output versus the input signal magnitude in dB (right).

Figures 4 and 5 illustrate the active and non-linear properties of the mammalian cochlea. Figure 4 depicts the traveling membrane-displacement wave propagating in the cochlea. The envelope of the wave exhibits peaking at the best place. The passive cochlear response (on the left) corresponds to either a dead cochlea as in the early experiments (Bekesy 1960), or in response to very loud incoming sounds, and consequently the peak is not highly tuned, and the amplification is not high. Later measurements performed on living cochleae exhibit much sharper frequency localization and much less damping for low sound levels (Geisler 1998) as in Figure 4 (right). This nonlinearity is further illustrated in Figure 5 showing that the response is highly tuned for quiet sounds below 30dB SPL, and the peak gain is up to 60dB. The cochlea exhibits essentially linear behavior in this region. For very loud sounds above 100dB SPL, the cochlear response is again linear and not very different from that of a dead cochlea. The peak is broad with a peak gain of about 0dB. However, within the range of normal acoustic input the cochlear response exhibits a strong compressive non-linearity at the peak. This is necessary for the auditory pathway to be able to resolve and interpret information encoded in varying sound frequencies and over wide range of sound levels, converting almost 120dB of input sound dynamic range into about 40dB of basilar membrane displacement.

As an amplifier and analyzer of sound, the cochlea acts as an active non-linear signal processor that performs its calculations in parallel, attaining an extremely wide dynamic range of 120dB over the wide frequency range that spans 3 decades with an extremely low noise and power consumption. The "cochlear amplifier" algorithm holds great promise to vastly improve the performance of the frequency analyzers operating over a very wide frequency range in low-power wide-dynamic-range applications.

1.1.2. Cochlear Implant Speech Processor Standard Architecture

Figure 6 shows an overview of a standard filter bank signal-processing chain in commercially available cochlear implant systems.

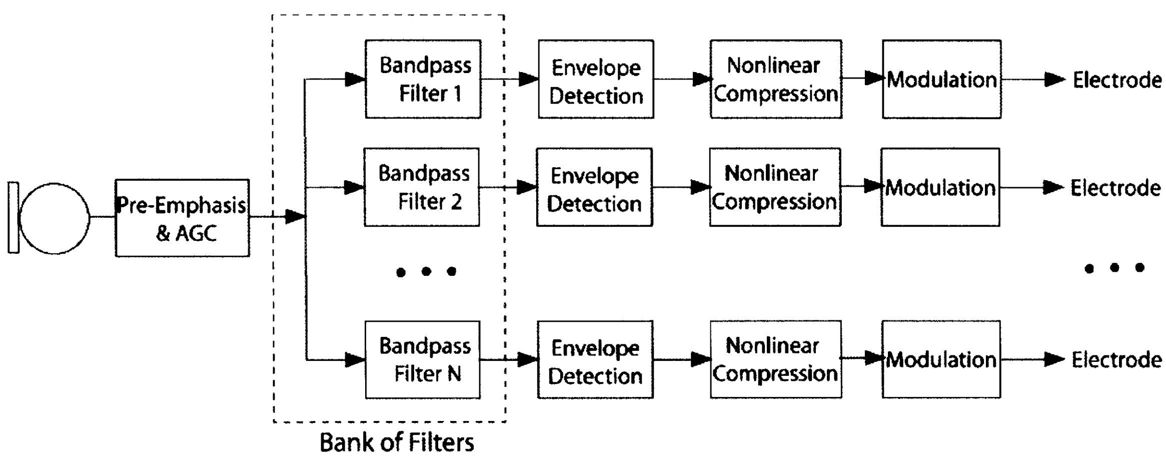


Figure 6: Standard filter bank architecture of the bionic ear.

The system mimics the function of the biological ear in stimulating neurons in the cochlea in response to sound. Only three channels of processing are shown although typical speech processors have 16 channels. Sound is first sensed by a microphone. Pre-emphasis filtering and automatic gain control (AGC) are then performed on the input. A bank of constant-Q wavelet-like bandpass filters decomposes the AGC output into different frequency bands. Envelope detectors then extract the envelope of the waveform in each channel. The dynamic range of each channel's envelope output is compressed to fit into the electrode dynamic range via the nonlinear compression blocks. Finally, a fixed-rate carrier is amplitude-modulated by the compressed envelope information and sent to the electrodes to create charge-balanced current stimulation (Loizou 1998).

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 now is to move to designs that can be fully implanted. Reducing the power of the speech processor is one of the keys to moving to a fully implanted system.

1.1.3. Cascaded Implementation of the Active Cochlea

If we want to construct a low-power, wide dynamic range frequency analyzer, using a system of distributed traveling-wave amplifiers is vastly more efficient than a bank of bandpass filters (Sarpeshkar 2000). Figure 7 shows a 117-stage 100Hz-to-10kHz cochlea that attains a dynamic range of 61dB while dissipating 0.5 mW implemented as an overlapping cascade of second-order low-noise lowpass filters (Sarpeshkar 1998).

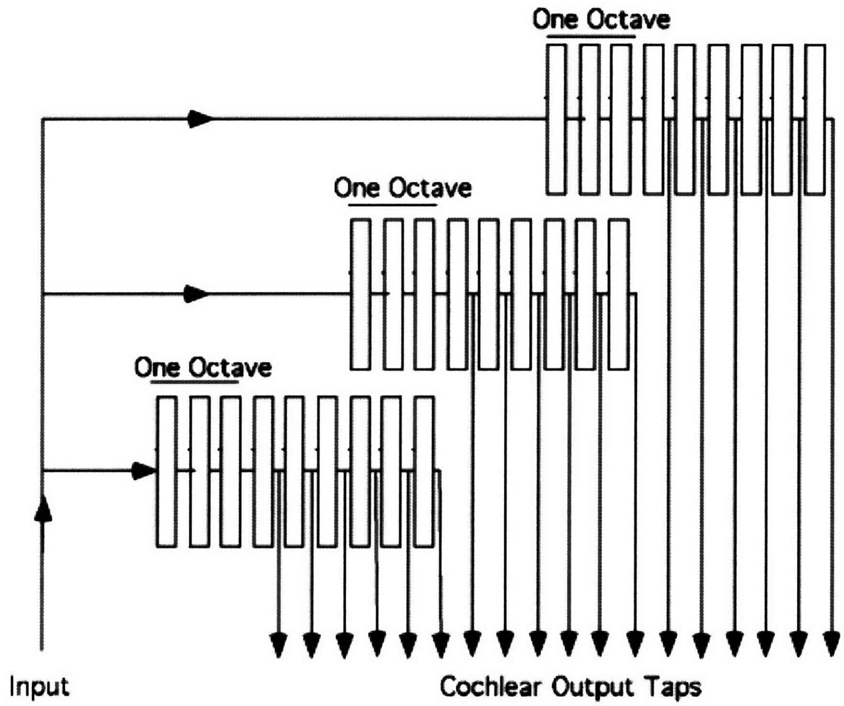


Figure 7: An overlapping cascade of the second-order lowpass filters, where the input is fed in parallel to smaller cochlear cascades whose corner frequencies overlap by 1 octave.

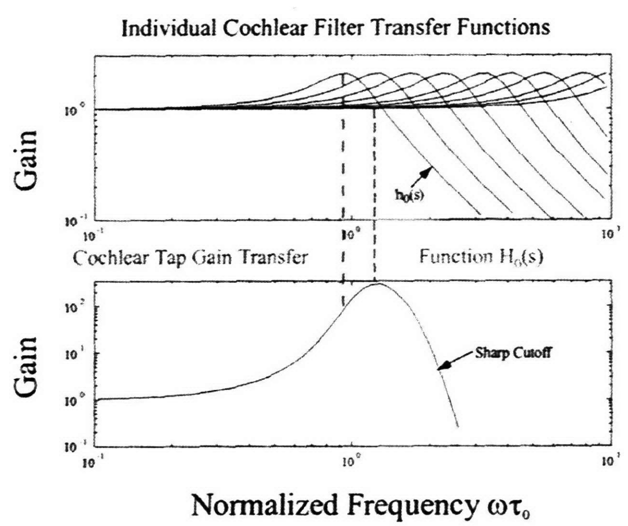


Figure 8: The cascade of lowpass second-order filters with low Q and exponentially tapered corner frequencies forms a bandpass transfer function with the high peak gain and sharp roll-off after the peak. Adapted from (Sarpeshkar 2000).

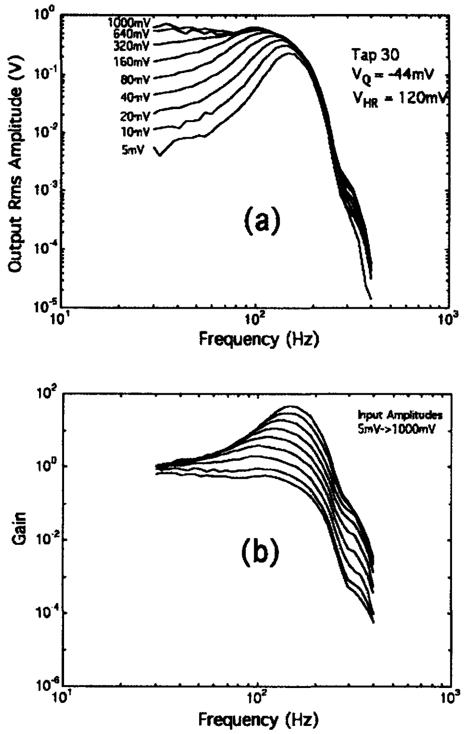


Figure 9: Frequency Response. (a) The frequency response for various input rms amplitudes is shown. Compare to Figure 5. (b) The corresponding gain of the cochlea (Sarpeshkar 1998). Compare to the experimental data from (Ruggero 1992).

This electronic cochlea faithfully reproduces many aspects of the biological cochlea. Figure 8 illustrates how the cascade of lowpass second-order filters with exponentially tapered corner frequencies forms a bandpass transfer function. Due to the distributed nature of the amplification in the cascade we can obtain a high peak gain and sharp roll-off after the peak even though the order and the Q of each individual filter are low.

Figure 9 demonstrates an experimentally measured bandpass frequency response of the electronic cochlea with maximum active amplification of about 50x at the peak, compressive nonlinearity at the peak for the normal input signals, and a sharp (10^{th} to 16^{th} order) roll-off after the peak.

However, this architecture has a range of issues like noise accumulation in the cascade, which is why the cochlear cascade was partitioned into an overlapping cascade structure. The group delay of the system was too high, which can be a problem in cochlear implant or speech recognition applications. In addition, the cochlear cascade was too compressive due to the local nature of the Q-adaptation. All these issues are inherent in the cascaded architecture, prompting the development of alternative approaches. One of the alternative approaches is modeling and implementation of the cochlea as a passive or active transmission-line-like structure.

1.1.4. Passive Cochlear Transmission-Line Model and its Implementation

The rectangular-box two-dimensional model of the cochlea is shown in Figure 10. The fluid is assumed to be incompressible, so that we can ignore the sound wave in the cochlear fluid, and consider only relatively slow traveling wave excitation.

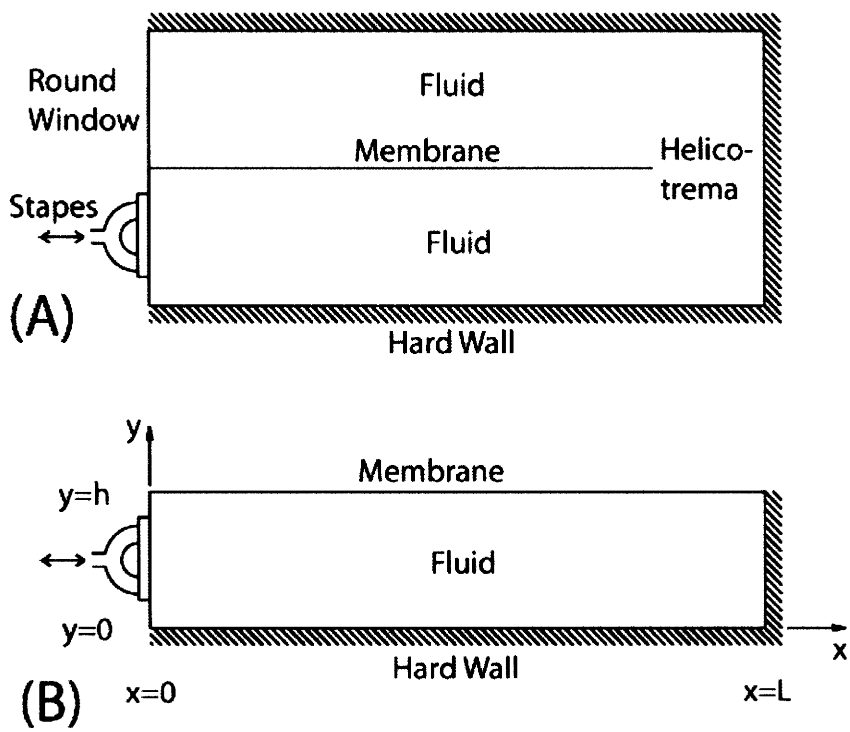


Figure 10: The physical two-dimensional model of the cochlea. (A) The model showing both chambers. (B) Fluid movement in both chambers assumed to be complementary in this approximation, so we can consider only one chamber. Adapted from (Watts 1993).

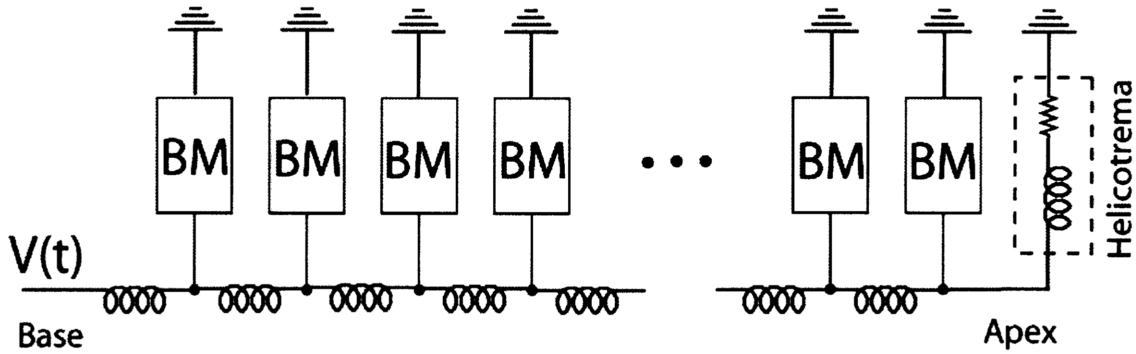


Figure 11: The electrical circuit equivalent of a one-dimensional single-line cochlear transmission-line model. In addition to the single-chamber approximation, the traveling wave is assumed to be much longer than the cross-sectional dimension of the cochlea. The inductances model the fluid mass and the boxes model the shunt admittance of the basilar membrane and organ of Corti, which vary along the length of the cochlea.

Figure 10 (A) shows both chambers of the cochlea with the basilar membrane and the organ of Corti in the center. Assuming that the basilar membrane with the organ of Corti move as a whole, the incompressible fluid displacement in both chambers is complementary, and we can consider only one chamber for the modeling, as in Figure 10 (B). If the voltage signals represent the pressure in the fluid, and the electric currents represent the volume velocities, we arrive at the electrical circuit equivalent of our model shown in Figure 11. In addition to the single-chamber approximation, the traveling wave is assumed to be much longer than the cross-sectional dimension of the cochlea. Therefore, there is no appreciable movement of the fluid in the y-direction, and the fluid mass can be modeled as one-dimensional array of inductors. The hydrodynamic impedance of the basilar membrane and organ of Corti, which vary along the length of the cochlea, is modeled by the set of electronic filters BM presenting the electric impedances $Z(j\omega,x)$. The series connection of the inductor and resistor models the mass and viscosity of the cochlear fluid moving through the small hole of the helicotrema. The motion of the stapes at the left side of the model drives the system and is represented by the input voltage *V(t)*. This model is referred to as the one-dimensional single-line cochlear transmission line. If we assume that the basilar membrane and organ of Corti present only acoustic compliance, viscosity and mass with no active processes inside, our model is passive and $Z(j\omega,x)=K(x)/j\omega+R(x)+j\omega\cdot M(x)$. Since the model is also linear time-invariant, we can divide the electrical impedances of all the elements by $j\omega$ for the ease of electronic implementation as shown in Figure 12. Now we need to implement $Z'(j\omega,x)=K(x)/(j\omega)^2+R(x)/j\omega+M(x)$ where masses become resistors, viscosities – capacitors, and the acoustic compliances become "supercapacitors".

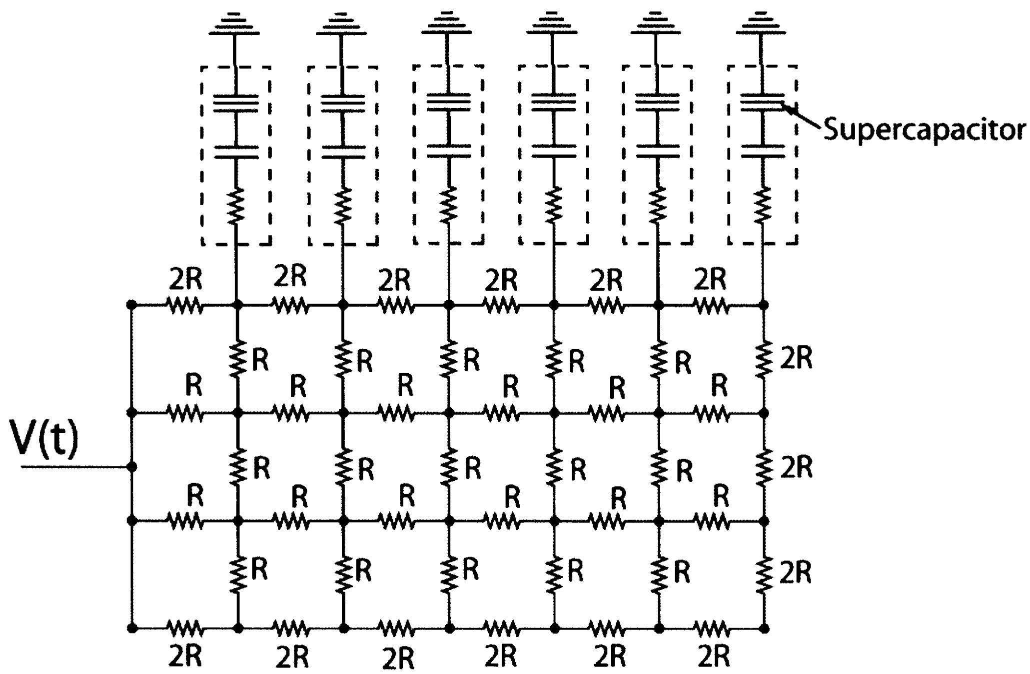


Figure 12: Two-dimensional cochlear circuit model. The chip has 64 stages, although we show only 6 for clarity. The resistive network models the cochlear fluid mass. The hard-wall boundary conditions are represented by the floating edges on the right and bottom sides of the network. The input signal V(t) is applied to the left end of the cochlea. The outputs are the currents flowing into the filter circuits at each stage.

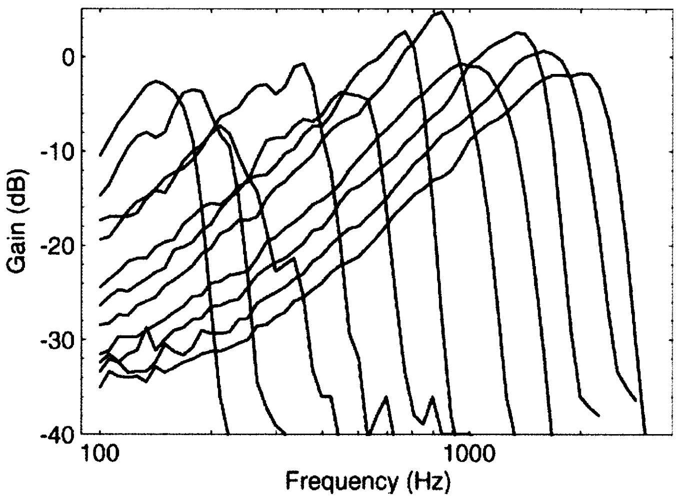


Figure 13: Experimentally measured magnitude of frequency response of every 5th current tap, from tap 10 to 60, in a 64-stage cochlear chip. Adapted from (Watts 1993).

A chip was fabricated in a CMOS process consisting of a two-dimensional 64x5 resistor array modeling the cochlear fluid mass, and 64 filter circuits modeling the passive impedance of the basilar membrane and organ of Corti. The increasing-mass scaling configuration, where $M(x) = M_0 e^{x/l}$, $R(x) = R_0$, $K(x) = K_0 e^{-x/l}$, was applied to the set of 64 filter circuits. The input signal V(t) is applied to the left end of the cochlea. The outputs are the currents flowing into the filter circuits at each stage. Figure 13 shows the frequency response magnitude of every 5^{th} current tap, from tap 10 to 60, measured on this cochlear chip.

This passive cochlear transmission-line model faithfully reproduces some aspects of the biological cochlea, for example a steep cut-off in the frequency response magnitude after the best place. But since this model assumes no active processes in the organ of Corti, it models a dead biological cochlea, and therefore lacks a very important feature of the real cochlear response – sharply tuned peak with a high gain near the best place. Active cochlear models attempt to solve this very important issue.

1.1.5. Active Cochlear Transmission-Line (TL) Models

1.1.5.1. One-dimensional single-line active cochlear TL model by (Zweig 1991).

One of the earliest and most successful attempts to build an active cochlear transmission-line model was to "derive" theoretically the hydrodynamic impedance of the basilar membrane and organ of Corti in Figure 11 based on the experimental data. Zweig assumed a certain type of scaling of the parameters along the length of the cochlea when they do not depend on the position x and frequency ω independently, but rather depend on their combination - a new independent variable $s_n \equiv j\omega \cdot e^{x/l} / \omega(0)$ (Zweig 1991). In fact, this is the same increasing-mass scaling configuration, where $M(x) = M_0 e^{x/l}$, $R(x) = R_0$, $K(x) = K_0 e^{-x/l}$, and the inductances in Figure 11 increase as

$\bar{M}(x) = \bar{M}_0 e^{x/l}$ with position too. He also assumed that the parameter values change insignificantly along the wavelength of the traveling wave. After confirming this assumption, Zweig employed the WKB approximation to solve the problem. He arrived at the following expression for the impedance of the basilar membrane and organ of Corti:

$Z(j\omega, x) \equiv Z(s_n) = M_0\omega(0)\cdot\left(s_n^2 + \delta s_n + 1 + \rho e^{-s_n\psi}\right)/s_n$, where

$$\omega^2(0) \equiv \frac{K_0}{M_0},\ \delta \equiv \frac{R_0}{M_0\omega(0)} = -0.1217,\ \rho = 0.1416,\ \psi = 1.742\cdot 2\pi.$$

Physically, the outer hair cells (OHCs) within the organ of Corti were presumed to provide both the active amplification resulting in negative damping δ, and the stabilizing term $\rho e^{-s_n\psi}$ that represents the compliance with pure delay. Although it remains unclear how the required acoustic impedance would be formed based on the anatomy of the organ of Corti, this model provides an excellent agreement with the available experimental data, faithfully reproducing both magnitude and phase of the frequency response, sharply tuned high gain peaks, steep roll-off after the best place, and even otoacoustic emissions. This model can also naturally incorporate cochlear nonlinearity by making the negative damping δ dependent on the signal level. This thesis intends to build upon this model. The only issue of practical implementation is that the pure delay $\rho e^{-s_n\psi}$ can not be built with a finite number of lumped elements in analog circuitry.

1.1.5.2. One-dimensional two-line active cochlear TL model by (Hubbard 1993).

Another active cochlear transmission-line model was developed by (Hubbard 1993). This model draws its inspiration from the traveling wave amplifier in RF design (Ginzton 1948). The traveling wave amplifier consists of an input transmission line, where the signal from the source propagates. The input line is tapped and the signal is coupled to a second transmission line via active elements. The signal in the second line experiences constructive interference from the multiple active devices and is amplified if the group velocities in both transmission lines are matched. In Hubbard's cochlear model shown in Figure 14, the input transmission line is replaced by the resonant passive-cochlea-like line, where the group velocity decreases exponentially as the traveling wave propagates along the line. The group velocity in the second line is chosen to be small, such that the group velocity match occurs at the best place of the first line. Significant amplification occurs here due to the active coupling.

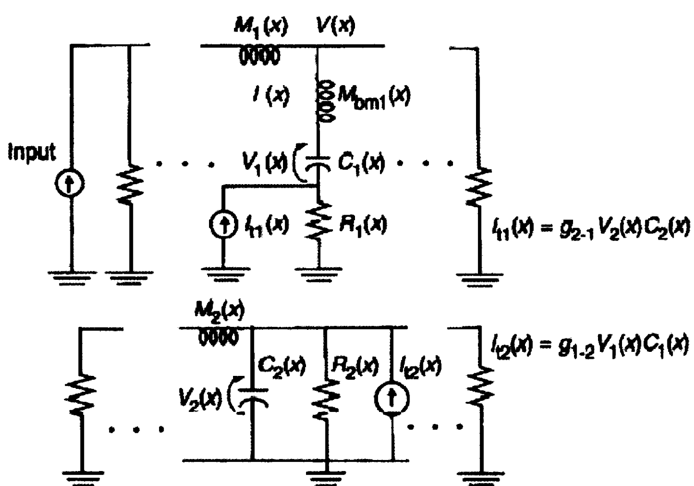


Figure 14: One section of the two-line one-dimensional active cochlear transmission-line model by (Hubbard 1993). Connecting 400 similar sections forms two coupled transmission lines, which are terminated at each end as shown.

The results from this model compare favorably with the experimental data, for example from the chinchilla (Ruggero 1990), as shown in Figure 15. Specifically, the height and the bandwidth of the peak response are in excellent agreement. One issue with this model is that its elements could not be mapped to the structures of the biological cochlea. This model is not friendly to an electronic implementation since it contains a lot of inductors, which would introduce a noise, complexity and power consumption hit associated with the audio-frequency electronic implementation of an inductor.

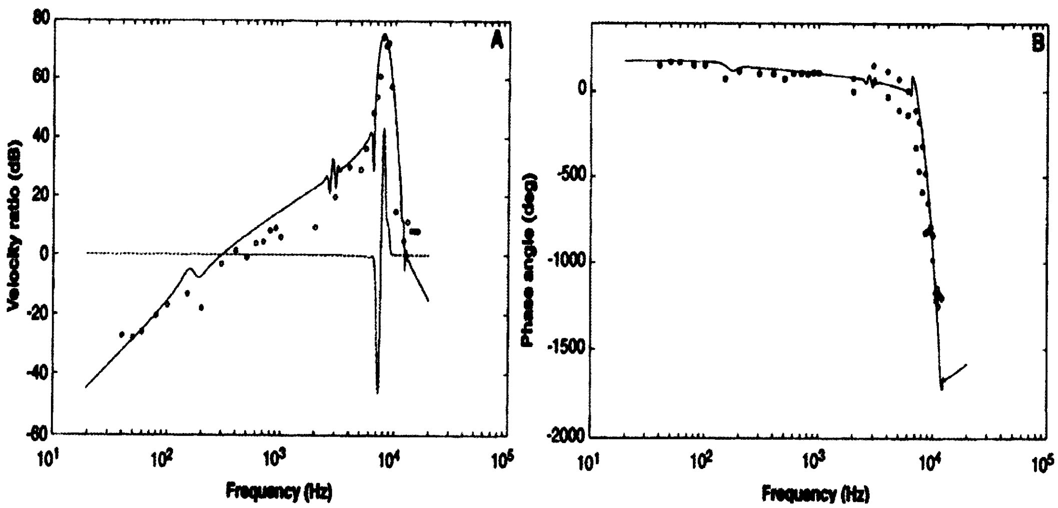


Figure 15: Comparison of Hubbard model data (solid lines) and experimental data (Ruggero 1990) from the chinchilla (dots). (A) The ratio of basilar membrane to stapes velocity. Calculated power transfer (dashed line) is also shown, on a linear scale. (B) Phase response. The points * and O are the extremes of the experimental data.

1.1.5.3. Active cochlear TL model by (Mammano 1993).

Mammano and Nobili proposed a model of the cochlea that can be described by the equation: $m \cdot \ddot{y} + r \cdot \dot{y} + k(x) \cdot y = p - \frac{\partial}{\partial x}\left(s \cdot \frac{\partial}{\partial x}\right)\dot{y}$, where y is the basilar membrane deflection, $r<0$ represents the net effect of the cochlear fluid viscous damping and the OHC's active force undamping action, m and $k(x)$ are mass and stiffness of the basilar membrane, and p is the pressure in the cochlear fluid that drives the basilar membrane and organ of Corti motion. The term $\frac{\partial}{\partial x}\left(s \cdot \frac{\partial}{\partial x}\right)\dot{y}$ describes the shearing motion between adjacent segments of the organ of Corti and s is the shearing resistance coefficient. This shearing motion provides a stabilizing action to the undamped cochlea, just like the pure delay term provided the stability in Zweig's cochlear model with negative resistance. A circuit representation of this model that corresponds to a one-dimensional (long-wave) approximation of the cochlear fluid motion is shown in Figure 16.

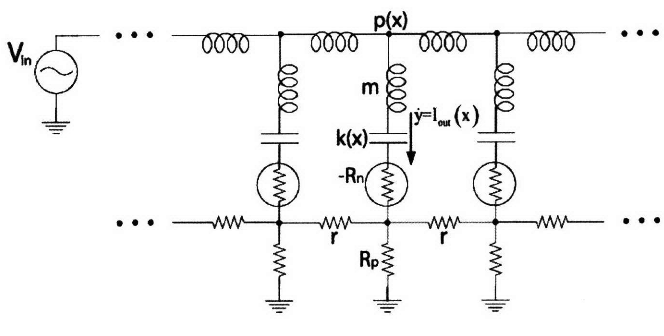


Figure 16: Circuit representation of a one-dimensional version of an active cochlear transmission-line model with negative and possibly nonlinear resistance *Rn* created by OHC action, and viscous stabilization with $s = \frac{R_p^2 \cdot \Delta x^2}{r}$. By (Mammano 1993).

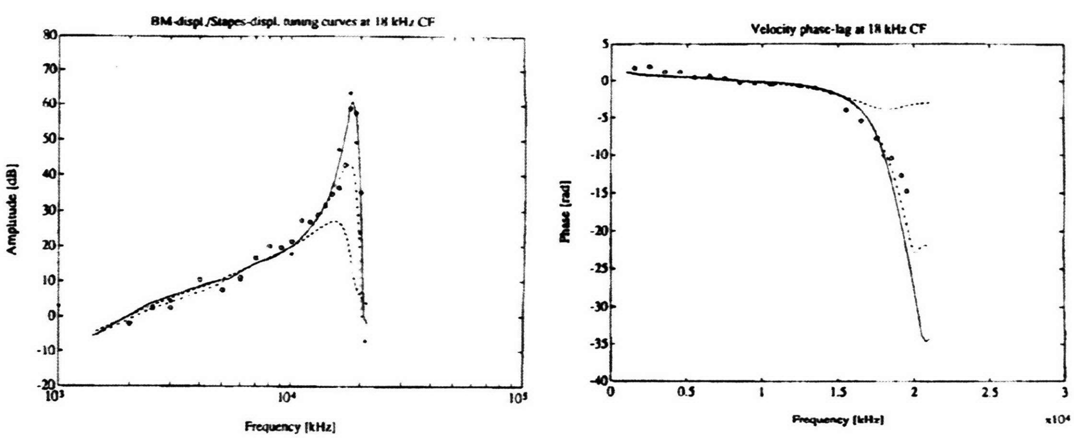


Figure 17: Basilar membrane velocity magnitude and phase response of the Mammano and Nobili model. Comparison with the experimental data of (Sellick 1982) (open circles).

The magnitude and phase of the frequency response of this model with the experimental data of (Sellick 1982) are shown in Figure 17. The major problem of this model is that it requires the OHC's undamping forces to exceed realistic value by about two orders of magnitude in order to produce the active amplification observed in experiment, as estimated in (Dimitriadis 1999). This casts doubt whether this model utilizes the collective action of the active elements efficiently, and would introduce excessive noise and power consumption should an electronic implementation be attempted.

1.1.5.4. Three-line active cochlear TL "Sandwich" model by (Hubbard 2000).

The biological three-compartment multi-mode wave-propagation model was proposed by Hubbard *et al.* in 2000. The interior of the Sandwich is the organ of Corti (OC), which is bounded by the reticular lamina (RL) with the fluid-filled scala vestibuli (SV) and the basilar membrane (BM) with the fluid-filled scala tympani (ST). Unlike all previous models, this model does not assume that the basilar membrane with the organ of Corti move as a whole, so the incompressible fluid displacement in both scala vestibuli and scala tympani is not assumed to be complementary. The circuit representation of this model that corresponds to a one-dimensional (long-wave) approximation of the cochlear fluid motion in all three compartments is shown in Figure 18. Hubbard *et al.* assumed the OHC active force production to be proportional to the OHC's stereocilia deflection, which is proportional to the displacement of the RL: $V_{active} = M \cdot \frac{I_{rl}(x)}{j\omega} \cdot e^{-x/l}$. Hubbard *et al.* was able to produce realistic results, shown in Figure 19, which utilized realistic OHC force production. Lu *et al.* took into account the slow time constant $\tau(x)$ due to the RC cutoff of the active potential in the OHC membrane: $V_{active} = M \cdot \frac{I_{rl}(x)}{j\omega \cdot (1 + j\omega \cdot \tau(x))} \cdot e^{-x/l}$.

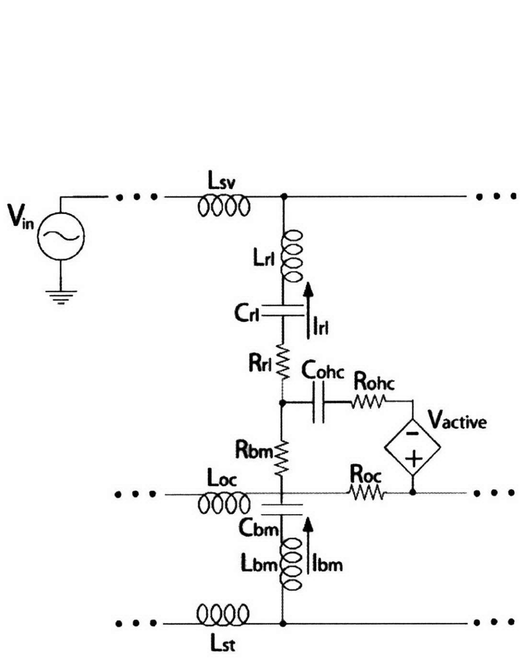


Figure 18: A circuit realization of an incremental section of the Sandwich model. $V_{active} = M \cdot \frac{I_{rl}(x)}{j\omega} \cdot e^{-x/l}$ form was used in (Hubbard 2000) model. A more realistic form $V_{active} = M \cdot \frac{I_{rl}(x)}{j\omega \cdot (1 + j\omega \cdot \tau(x))} \cdot e^{-x/l}$ was used in the (Lu 2005) model to account for the slow OHC membrane time constant $\tau(x)$. M can be varied to study the effects of nonlinearity.

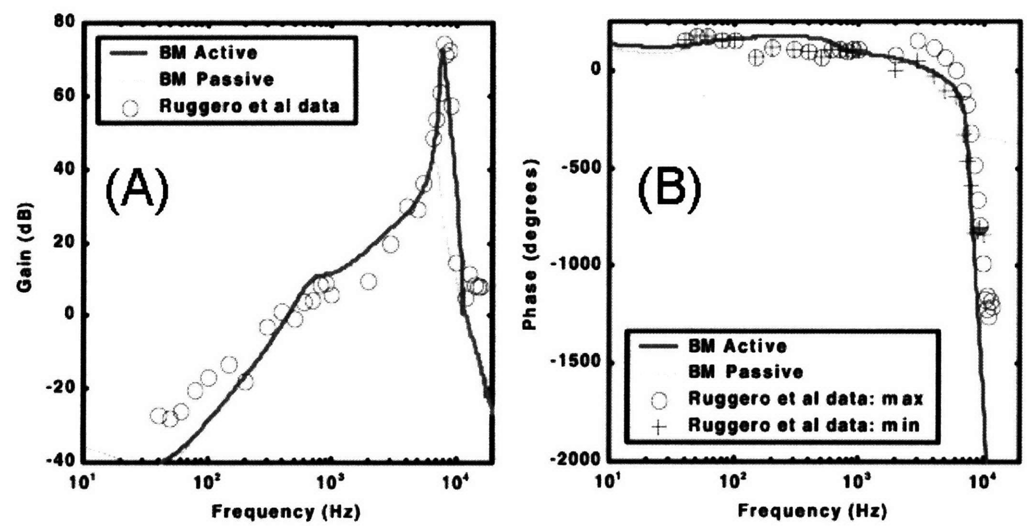


Figure 19: Comparison of (Hubbard 2000) Sandwich model data (solid lines - active, and dashed lines – passive responses) and experimental data (Ruggero 1990) from the chinchilla (dots). (A) The ratio of basilar membrane to stapes velocity. (B) Phase angle of responses relative to stapes velocity. The points + and O are the extremes of the experimental data.

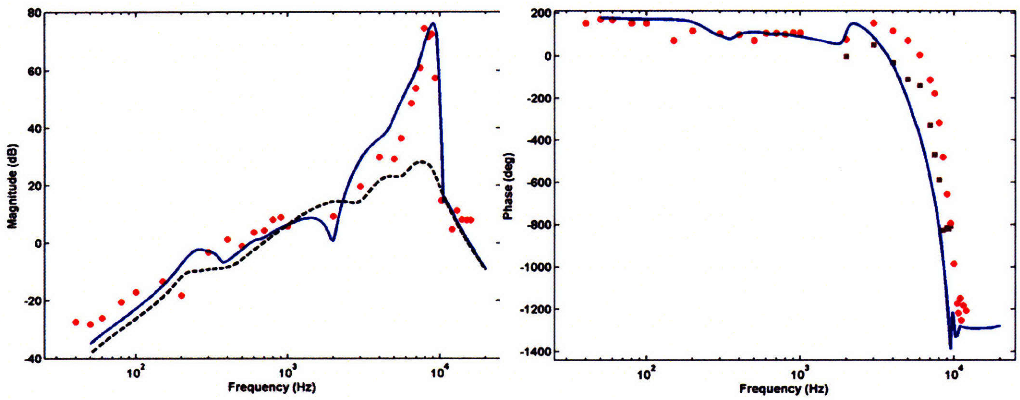


Figure 20: Sandwich model with realistic OHC active force taking the slow OHC membrane time constant into account by (Lu 2005). Comparison of the results (solid lines - active, and dashed lines – passive responses) and experimental data (Ruggero 1990) from the chinchilla (dots). The ratio of basilar membrane to stapes velocity (left); phase (right).

The data produced with this more realistic form of the OHC active force is shown in Figure 20. Thus, this result explains how a slow OHC with realistic force production enables fast cochlear amplification via a negative feedback mechanism (Lu 2005). The effect of nonlinearity can also be approximated by varying the level of the OHC active force production M. By setting $M=0$, a passive cochlear response was obtained, shown in Figure 20. More extensive research on the effects of the nonlinearities is planned. While this model is an excellent candidate for a parameter-tolerant biological cochlear amplifier, it is not friendly to an electronic implementation. Preliminary results show that the OC (third) line is not essential to reproducing the cochlear features faithfully. But even a reduced two-line model contains a lot of inductors, which would introduce a noise, complexity and power consumption hit associated with the audio-frequency inductor implementation.

1.2. Analog Filter Topologies: Gm-C and Log-Domain Topologies

Two classes of topologies have emerged in analog filtering applications: Gm-C and log-domain. The Gm-C topology is defined as filters built using linear voltage to current converters (Gm) and capacitors (C). Sanchez-Sinencio and Silva-Martinez provided an excellent overview of Gm-C filters (Sanchez-Sinencio 2000). The log-domain topology, also known as translinear, current-mode, or companding filters exhibit in theory an externally linear frequency-dependent transfer function even though the internal signal path contains nonlinear elements. An excellent general overview of companding filters can be found in (Tsividis 1997).

1.2.1. G_m-C topology overview.

The simplest implementation of the most common differential transconductor is shown in Figure 21.

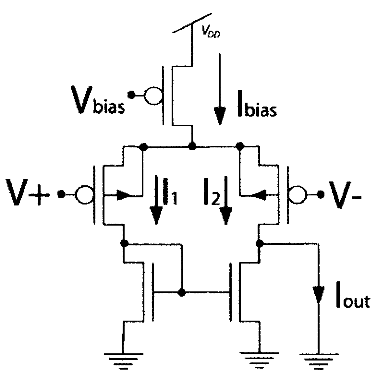


Figure 21: The simplest 5-transistor OTA transconductor.

The differential pair splits the current I_{bias} between two legs. The current mirror formed by the bottom two transistors performs the current subtraction to form I_{out}. Using the source-referenced transistor model in the subthreshold region, we can derive I_{out}:

$$I_1 = I_s e^{-\frac{\kappa(V_+ - V_s)}{\phi_t}} \quad ; \quad I_2 = I_s e^{-\frac{\kappa(V_- - V_s)}{\phi_t}}$$

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

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

$$I_{out} = I_{bias}\left(\frac{e^{-\frac{\kappa V_-}{\phi_t}} - e^{-\frac{\kappa V_+}{\phi_t}}}{e^{-\frac{\kappa V_+}{\phi_t}} + e^{-\frac{\kappa V_-}{\phi_t}}}\right) \quad \Rightarrow \quad I_{out} = I_{bias}\tanh\left(\frac{V_+ - V_-}{2\phi_t / \kappa}\right)$$

The linear range is $2\phi_t / \kappa \approx 75mV$. If the input voltage swing $V_+ - V_-$ is below that value, the output current is assumed to be approximately linear with the input voltage, and the linearization gives $G_m = \dfrac{I_{bias}}{2\phi_t / \kappa}$.

Linear range is one of the major limitations in low-power wide-dynamic-range applications. A wide variety of techniques have been used to improve this linear range, but they can be broadly divided into three categories: attenuation, degeneration, and nonlinear term cancellation (Sarpeshkar 1997).

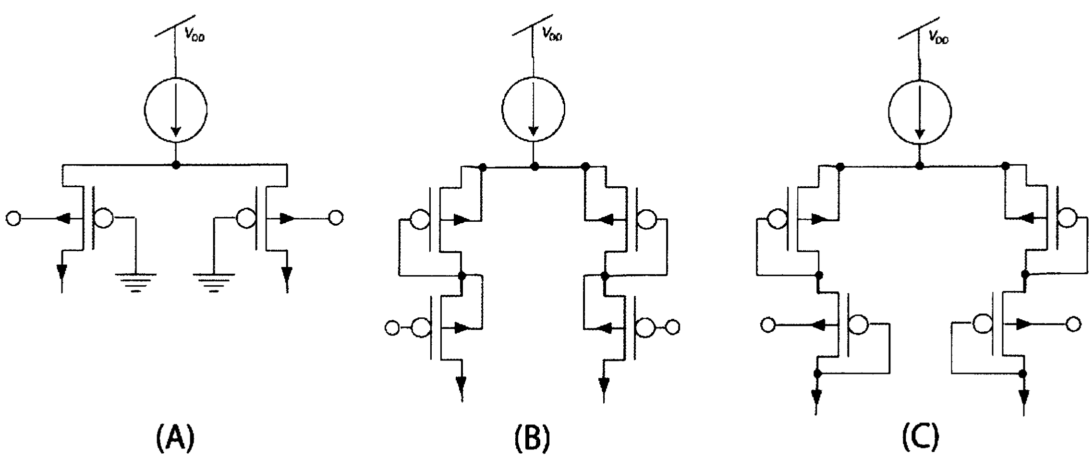


Figure 22: Linear range enhancement techniques. (A) Well attenuation. (B) Diode degeneration. (C) A combination of well attenuation, and diode and gate degeneration.

Attenuation is the simplest of the techniques; the signal is simply scaled by a factor less than 1 prior to controlling the differential pair. By using the well as the input to the circuit rather than the gate, as shown in Figure 22 (A), the transconductance of the differential pair is decreased. Degeneration schemes also lower the voltage across the control terminal, but they do it through feedback. The circuit in Figure 22 (B) shows 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 the transconducting device 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 of the DC voltage drop across the diodes. Figure 22 (C) shows a combination of well attenuation, diode degeneration, and gate degeneration. The voltage induced across diode-connected transistors lowers the voltage difference between the source and the gate. Both terminals act to lower the current (Sarpeshkar 1997).

The technology scaling and low-power applications require decreased supply voltages. But mixed-signal applications require a relatively large threshold voltage V_T to limit transistor OFF currents in the digital part. This makes the analog design difficult as V_T becomes a very large part of the supply voltage. While the techniques shown in Figure 22 (B) and (C) are quite effective in enhancing the linear range of the transconductor, they require larger *Vdd*. One of the ultra-low power-supply-voltage techniques was developed by Chatterjee, Kinget and Tsividis in 2004 (Chatterjee 2004). A fully-differential gain stage running on a 0.5V power-supply with local common-mode feedback is shown in Figure 23. The V_T of the devices was about 0.5V.

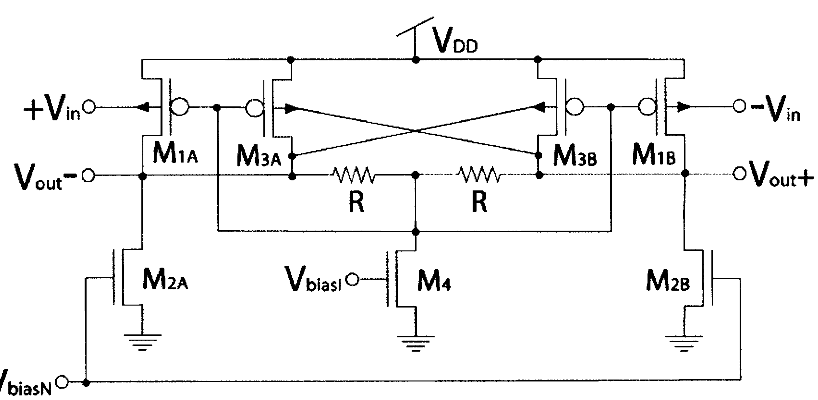


Figure 23: A 0.5V fully differential gain stage. The V_T of the devices is about 0.5V.

In this circuit V_{biasN} biases the input transconducting devices M_{1A} and M_{1B}. Resistors R detect the output common mode voltage which is fed back to the gates of M_{1A}, M_{3A}, M_{1B}

and M_{3B} for common mode rejection. The output common mode DC voltage is set by pulling a small current through resistors R with V_{biasI}. The bulk-inputs of M_{3A} and M_{3B} form a cross coupled pair that adds an incremental negative resistance to the output, which boosts the differential gain and positive resistance, further decreasing the common mode gain (Chatterjee 2004).

1.2.2. Log-domain topology overview.

Unlike in Gm-C, log-domain filters do not approximate the transconductor as a linear element. Instead, the transistor's exponential relationship between the input voltage and the output current is exploited.

We proceed to describe the most powerful log-domain circuit synthesis technique (Frey 1996). Suppose that we have the following system of $N=2$ equations:

$$\dot{x}_1 = A\cdot x_1 + B\cdot x_2 + E\cdot u$$
$$\dot{x}_2 = C\cdot x_1 + D\cdot x_2 + F\cdot u$$

This state-space representation implements the transfer function of the order N:

$$\frac{x_1}{u} = \frac{s\cdot E + (B\cdot F - D\cdot E)}{s^2 - s\cdot(A+D) + (A\cdot D - B\cdot C)};$$
$$\frac{x_2}{u} = \frac{s\cdot F + (C\cdot E - A\cdot F)}{s^2 - s\cdot(A+D) + (A\cdot D - B\cdot C)}.$$

It is obvious that there is some freedom in choosing A, B, C, D, E and F in practice. The implementation of the Nth order transfer function requires N state equations with N state variables, thus N capacitors are needed. Applying the exponential mapping:

$$x_i = I_i \cdot e^{V_i/U_t};\quad u = I_u \cdot e^{V_{in}/U_t}$$

Where I_i, I_u - some DC currents, V_i - ith capacitor voltage and U_t - a constant that equals ϕ_t for bipolar, and ϕ_t/κ for subthreshold CMOS implementations. The ith capacitor's current is:

$$I_{capi} \equiv C_i \cdot \dot{V}_i = C_i \cdot U_t \cdot \frac{\dot{x}_i}{x_i}$$

Denoting the DC, but not necessarily positive currents with the following equations,

$$I_A = C_1 \cdot U_t \cdot A, \quad I_B = C_1 \cdot U_t \cdot B \cdot \frac{I_2}{I_1}, \quad I_E = C_1 \cdot U_t \cdot E \cdot \frac{I_u}{I_1},$$
$$I_C = C_2 \cdot U_t \cdot C \cdot \frac{I_1}{I_2}, \quad I_D = C_2 \cdot U_t \cdot D, \quad I_F = C_2 \cdot U_t \cdot F \cdot \frac{I_u}{I_2}$$

Our state-space representation becomes:

$$I_{cap1} = I_A + I_B \cdot e^{(V_2 - V_1)/U_t} + I_E \cdot e^{(V_{in} - V_1)/U_t}$$
$$I_{cap2} = I_C \cdot e^{(V_1 - V_2)/U_t} + I_D + I_F \cdot e^{(V_{in} - V_2)/U_t}$$

Note that the components of each I_{capi} should be of different signs. For example, I_A, I_B, and I_E cannot all be positive, otherwise I_{cap1} is always positive and the capacitor voltage can only increase. That circuit will not work. This condition imposes some limitation on our freedom in choosing *A*, *B*, *C*, *D*, *E* and *F*. Sometimes this limitation is so severe that no constants can be chosen to satisfy it and implement the required transfer function simultaneously. In this case the operating point can be adjusted by adding an additional input to the state space representation. Because the filter is externally linear, a DC value applied to this input will simply shift the output.

In order to implement any state-space representation, we just need to implement

$$I_{cap} = \pm I_0 \cdot e^{(V_0 - V_{cap})/U_t}, \quad I_0 > 0$$

Some of the building block circuits used by Frey are shown in Figure 24 (Frey 1996).

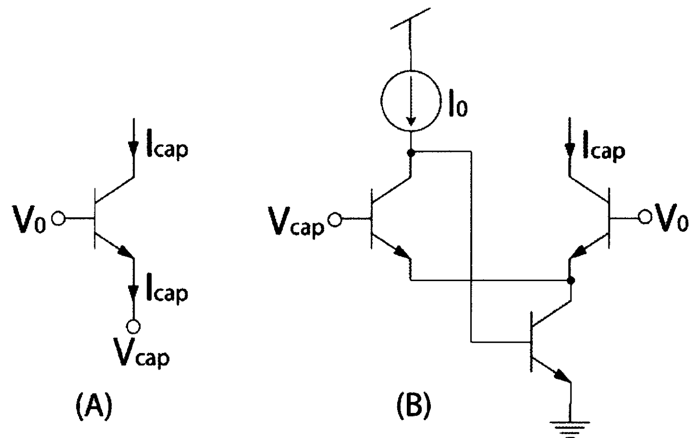


Figure 24: Circuits implementing: (A) $I_{cap} = +I_s \cdot e^{(V_0 - V_{cap})/U_t}$ (B) $I_{cap} = -I_0 \cdot e^{(V_0 - V_{cap})/U_t}$

The Dynamic Trans Linear (DTL) is another current-mode circuit synthesis technique. Consider the basic building block shown in Figure 25 (Mulder 2001).

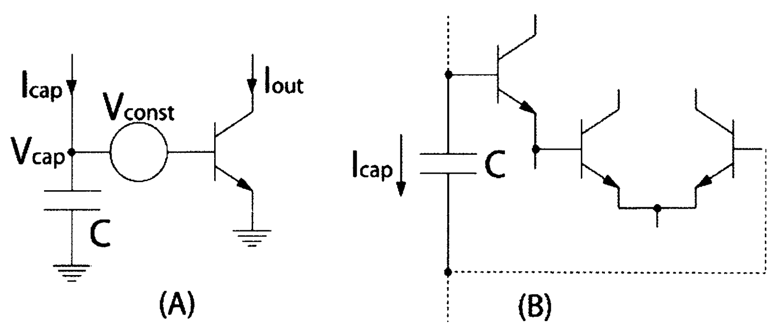


Figure 25: Principle of DTL circuits: (A) basic building block; (B) DTL loop.

The basic building block shown in Figure 25 (A) can be simply analyzed.

$$I_{out} = I_s \cdot e^{V_{be}/U_t} \quad \Rightarrow \quad \frac{\dot{I}_{out}}{I_{out}} = \frac{\dot{V}_{be}}{U_t} = \frac{\dot{V}_{cap}}{U_t}$$

$$I_{cap} = C \cdot \dot{V}_{cap} = CU_t \frac{\dot{I}_{out}}{I_{out}} \quad \Rightarrow \quad CU_t \cdot \dot{I}_{out} = I_{cap} \cdot I_{out}$$

The last equation states the DTL principle: A time derivative of a current can be mapped onto a product of currents. And the product of currents can be realized using Gilbert multipliers, allowing for the implementation of a linear or nonlinear differential equation.

Figure 25 (B) shows a generalization of the DTL principle. A corresponding equation is:

$$I_{cap} = CU_t \cdot \sum_i \pm \frac{\dot{I}_{out,i}}{I_{out,i}}$$

The ± sign of each term depends on the orientation of the corresponding transistor. This equation is applied to each capacitance in the circuit. Elimination of the intermediate currents yields the differential equation describing the output current (Mulder 2001).

Log-domain circuits can be operated in a class-AB mode to improve performance (Frey 1999, Serdijn 1999). The log-domain circuits as presented require that I_{in} include a DC offset such that it never becomes negative. This DC current is equal to half of the maximum signal swing. As the signal grows it begins to clip on the bottom as shown in Figure 26 (b). But on the top there is no clipping even if the signal amplitude is much larger than the offset current.

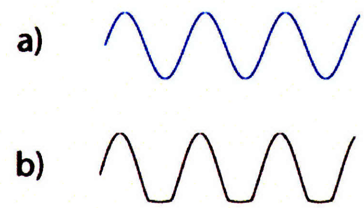


Figure 26: Clipping.

This property of log-domain filters allows the creation of a special type of differential circuits. 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 27 demonstrates the difference.

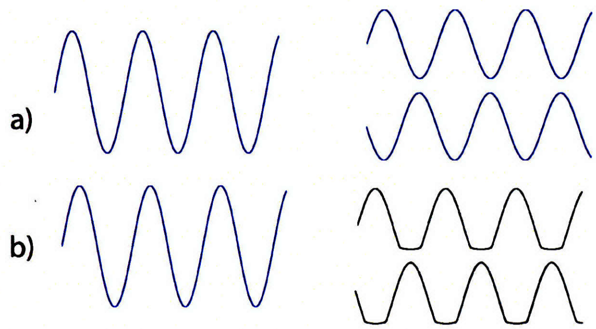


Figure 27: Differential signals: a) constant common mode; b) class AB.

Current splitting is a nonlinear process. Practical nonlinear devices present a dead-zone that should be overcome. To do so consumes power and limits the dynamic range of the filter. A low-power wide-dynamic-range current splitter is presented in (Zhak 2003).

The noise that log-domain class-AB circuits produce has very different properties from that of the G_m-C circuits. In G_m-C circuits, where all transfer functions are linear and the noise sources are stationery, the output noise is independent of the signal level. On the contrary, in log-domain class-AB circuits' noise depends on the signal level in a way that the signal-to-noise ratio stays approximately constant (Serdijn 1999). Heuristically, G_m-C filters correspond to fixed-point arithmetic signal processing, whereas log-domain class-AB circuits behave more like floating-point arithmetic. Log-domain class-AB circuits also provide other benefits like small voltage excursions.

1.3. Thesis outline

This thesis is organized as follows. In Chapter 2 we present the design of the novel low-power wide dynamic range envelope detector, which was developed to implement a standard cochlear implant speech processor. 60-dB dynamic range version was used in a channel, and 75-dB 1-μA version was used for the input automatic gain control of that chip. This envelope detector is one of the most important part of our cochlear implementation, providing both gain adjustment of the section to mimic biological cochlea, and adjusting the bias currents of the filter in the section implementation to lower noise at the low signal levels and lower the distortion at the high signal levels. In Chapter 3 we present our two novel cochlear architectures. Both are useful in implementation of the cochlear algorithm. We choose to realize the cascade of all-pass second-order filters for our audio-frequencies application as this implementation provides the same benefits as active transmission-line cochlear model reducing the phase lag and the group delay in the cascade, sharpening the high-frequency slope and increasing Q_{-10dB}, and improving nonlinear compression characteristics of the system. In Chapter 4 we present novel technique for analyzing multi-mode transmission-line cochlear models. Chapter 5 presents a novel design method for very efficient implementation of high-Q log-domain filters, which our architecture requires. Combined with adjustment of biases in those filters, our method enables cutting power consumption by a factor of Q and simultaneous improvement in maximum SNR by Q and extension of the dynamic range by a factor of Q^4. Chapter 6 presents experimental data from our electronic cochlea implementation. Chapter 7 summarizes our work and suggests directions for future improvement and commercialization of our system.

1.4. References

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

Bekesy, G. von **(1960)** "Experiments in hearing," McGraw Hill, NY.

Chatterjee, S., Tsividis, Y., and Kinget, P. **(2004)** "A 0.5 V Bulk-Input Fully Differential Operational Transconductance Amplifier," *Proc. European Solid-State Circuits Conference.*

Dallos, P., Fakler, B. (**2002**) "Prestin, a New Type of Motor Protein," *Nature Reviews Molecular Cell Biology,* **3**, pp. 104-111.

Dimitriadis, E. K., Chadwick, R. S. (**1999**) "Solution of the inverse problem for a linear cochlear model: A tonotopic cochlear amplifier," *J. Acoust. Soc. Am.* **106**, 1880-1892.

Frey, D. R. (**1996**) "Log Domain Filtering for RF Applications," *IEEE Journal of Solid-State Circuits*, **31** (10), pp. 1468-1475.

Frey, D. R., and Tola, D. R. (**1999**) "A State-Space Formulation for Externally Linear Class AB Dynamical Circuits," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing,* **46** (3), pp. 306-314.

Geisler, C. (**1998**) "From Sound to Synapse: Physiology of the Mammalian Ear," New York: Oxford University Press.

Ginzton, E., Hewlett, W., Jasberg, J., Noe, J. (**1948**) *Proc. IRE* **36**, p. 956.

Hubbard, A. (**1993**) "A Traveling-Wave Amplifier Model of the Cochlea," *Science*, **259**, pp. 68-71.

Hubbard, A. E., Yang, Z., Shatz, L., and Mountain, D. C. (**2000**) "Multi-Mode Cochlear Models," *Proceedings of the International Symposium on Recent Developments in Auditory Mechanics, (2000),* pp. 167-173.

Kessel, R. G., Kardon, R. H. (**1979**) "Tissues and Organs: A Text-Atlas of Scanning Electron Microscopy," W. H. Freeman, New York.

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

Lu, T., S. Zhak, P. Dallos, and R. Sarpeshkar (**2005**) "A Model for Fast Cochlear Amplification with Slow Outer Hair Cells," *Proceedings of the International Symposium on Auditory Mechanisms: Processes and Models*, Portland, Oregon, July 23-28, 2005.

Mammano, F., Nobili, R. (**1993**) "Biophysics of the cochlea: Linear approximation," *J. Acoust. Soc. Am.* **93** (6), pp. 3320-3332.

Mulder, J., Serdijn, W. A., A. C. van der Woerd, and A. H. M. van Roermund (**2001**) "A Generalized Class of Dynamic Translinear Circuits," *IEEE Transactions on Circuits and Systems II: Analog and Digital Signal Processing*, **48** (5), May, pp. 501-504.

Ruggero, M., Rich, N., Robles, L., Shivapuja, B. (**1990**) *J. Acoust. Soc. Am.* **87**, p. 1612.

Ruggero, M. A., Rich, N. C., and Recio, A. (**1992**) "Basilar membrane responses to clicks," *in Auditory Physiology and Perception*, Pergamon Press, London, pp. 85-91.

Sanchez-Sinencio, E., and Silva-Martinez, J. (**2000**) "CMOS Transconductance Amplifiers, Architectures and Active Filters: a Tutorial," *IEE Proc.-Circuits Devices Syst.*, **147** (1), pp. 3-12.

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

Sarpeshkar, R., Lyon, R., Mead, C. **(1998)** "A Low-Power Wide-Dynamic-Range Analog VLSI Cochlea," *Analog Integrated Circuits and Signal Processing*, **16**, pp. 245-274.

Sarpeshkar, R. **(2000)** "Traveling waves versus bandpass filters: The silicon and biological cochlea," *Proceedings of the International Symposium on Recent Developments in Auditory Mechanics, (2000),* pp. 216-222.

van Schaik, A., Fragniere, E. **(2001)** "Pseudo-Voltage-Domain Implementation of a Design of a 2-Dimensional Silicon Cochlea," *Proceedings of the IEEE International Symposium on Circuits and Systems (ISCAS 2001),* **3**, pp. 185-188.

Sellick, P. M., Patuzzi, R., and Johnstone, B. M. **(1982)** "Measurement of basilar membrane motion in the guinea pig using the Mossbauer technique," *J. Acoust. Soc. Am.* **72**, 131-141.

Serdijn, W. A., Kouwenhoven, M. H. L., Mulder, J., and A. H. M. van Roermund **(1999)** "Design of High Dynamic Range Fully Integratable Translinear Filters," *Analog Integrated Circuits and Signal Processing*, **19**, pp. 223-239.

Spelman, F. A. **(1999)** "The Past, Present, and Future of Cochlear Prostheses," *IEEE Engineering in Medicine and Biology*, p. 27.

Tsividis, Y. **(1997)** "Externally linear, time-invariant systems and their application to companding signal processors," *IEEE Tran. on Circuits and Systems II: Analog and Digital Signal Processing*, **44** (2), pp. 65 – 85.

Watts, L. **(1993)** "Cochlear Mechanics: Analysis and Analog VLSI," *Ph.D. Thesis*, Caltech.

Zhak, S. M., Baker, M. W., and Sarpeshkar, R. **(2003)** "A Low-Power Wide Dynamic Range Envelope Detector," *IEEE Journal of Solid-State Circuits*, **38** (10), pp. 1750-1753.

Zhak, S. M., Mandal, S., and Sarpeshkar, R. **(2004)** "A Proposal for an RF Cochlea," *in Proc. Asia Pacific Microwave Conference.*

Zweig, G. **(1991)** "Finding the impedance of the organ of Corti," *J. Acoust. Soc. Am.*, **89**, pp. 1229-1254.

2. A Low Power Wide Dynamic Range Envelope Detector

***Abstract*—We report a 75dB, 2.8µW, 100Hz-10kHz envelope detector in a 1.5µm 2.8V CMOS technology. The envelope detector performs input-dc-insensitive voltage-to-current-converting rectification followed by novel nanopower current-mode peak detection. The use of a subthreshold wide-linear-range transconductor (WLR OTA) allows greater than 1.7Vpp input voltage swings. We show theoretically that this optimal performance is technology-independent for the given topology and may be improved only by spending more power. A novel circuit topology is used to perform 140nW peak detection with controllable attack and release time constants. The lower limits of envelope detection are determined by the more dominant of two effects: The first effect is caused by the inability of amplified high-frequency signals to exceed the deadzone created by exponential nonlinearities in the rectifier. The second effect is due to an output current caused by thermal noise rectification. We demonstrate good agreement of experimentally measured results with theory. The envelope detector is useful in low power bionic implants for the deaf, hearing aids, and speech-recognition front ends. Extension of the envelope detector to higher-frequency applications is straightforward if power consumption is increased.**

***Index Terms*—Bionic Ear, Cochlear Implant, Envelope Detector, Rectifier, Peak Detector, Ultra-Low Power, Hearing Aids**

2.1. Introduction

BIONIC ears (BE's) 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 1 shows an overview of a common signal-processing chain. Only four channels of processing are shown although typical BE's have 16 channels. Sound is first sensed by a microphone. Pre-emphasis filtering and automatic gain control (AGC) are then performed on the input. Analog implementations of the AGC require envelope detection to be performed [2]. Bandpass filters (BPF's) divide the AGC output into different frequency bands. Envelope Detectors (ED's) then detect the envelope of the waveform in each channel. The dynamic range

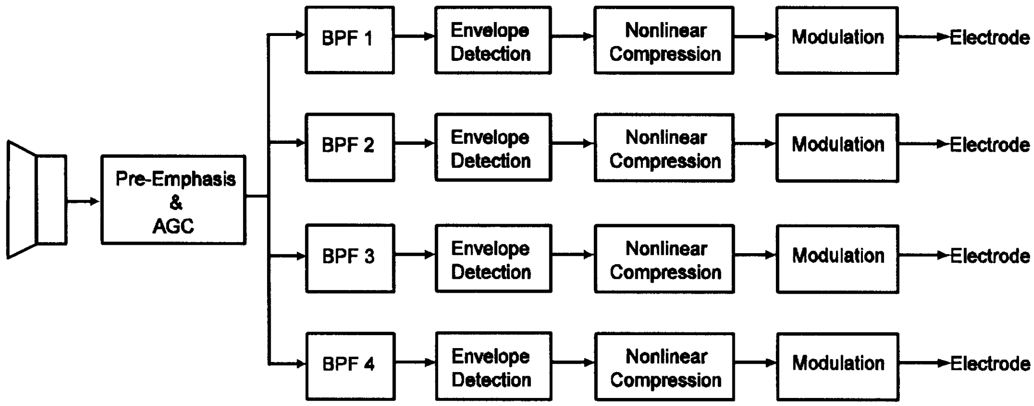


Figure 1: Bionic Ear Overview.

of each channel's envelope output is compressed to fit into the electrode dynamic range via the nonlinear compression blocks. Finally, the signals from each channel are modulated by the compressed envelope information and sent to the electrodes to create charge-balanced current stimulation [3].

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 now is to 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 all-analog processing strategies promise power reductions of an order of magnitude over even advanced DSP designs [4, 5, 6, 7]. We would like to implement envelope detectors with microwatt and submicrowatt power consumption to serve as building blocks in such ultra low power all-analog processing implementations.

Portable speech-recognition systems of the future will likely have more analog processing before digitization to reduce the computational bandwidth on the DSP and save power. The front end for such systems is remarkably similar to that shown in Figure 1 for bionic ear processing. Envelope detection is required for gain control and spectral energy estimation. Hearing aids perform broadband and multiband compression and require envelope detection for gain control and spectral energy estimation as well. Since the input to our envelope detector is a voltage but the output of the envelope detector is a current, translinear circuits can be used to implement a wide range of nonlinear functions on the output currents, which is useful for compression [8]. Thus, the envelope detector that we discuss in this paper is likely to have wide applicability in audio applications like implant processing, speech recognition, hearing aids.

If one is willing to increase power consumption, extensions to higher frequency applications like sonar or RF demodulation appear straightforward although we have not investigated the use of the envelope detector in such applications. Throughout the rest of the paper, we shall focus on the bionic ear application since that is the primary motivation for this work.

The BE application offers a number of constraints on the design of envelope detectors. It is battery powered and required to run off a low voltage; this design is optimized for 2.8 Volts. The envelope detector must provide frequency-independent operation over most of the audio range, from 100Hz to 10kHz. It should have a dynamic range of at least 60dB for narrowband envelope detection, and 70dB for broadband envelope detection. It must be insensitive to the input DC voltage providing a DC-offset-free current output. The envelope detector should have an adjustable attack time constant of around 10ms, and an adjustable release time constant of around 100ms. And, most importantly, it must minimize power while achieving all these specifications.

The organization of this chapter is as follows. In Section 2.2, we discuss the design of the voltage-to-current-converting rectifier, the first half of the envelope detector. In Section 2.3, we discuss the design of the current-mode peak detector, the second half of the envelope detector. In Section 2.4 we present experimental results from a chip. Finally, in Section 2.5, we conclude by summarizing the key contributions.

2.2. Rectifier Design

The basic current-converting rectifier topology examined here is a subthreshold G_m-C first-order high-pass filter, where the current through the capacitor is split into a positive half and a negative half by an intervening class-B mirror. Figure 2 shows the circuit. We can use one or both halves of the current in the rectifier output depending on whether we wish to perform half-wave or full-wave rectification respectively. Circuit operation is based on the fact that provided,

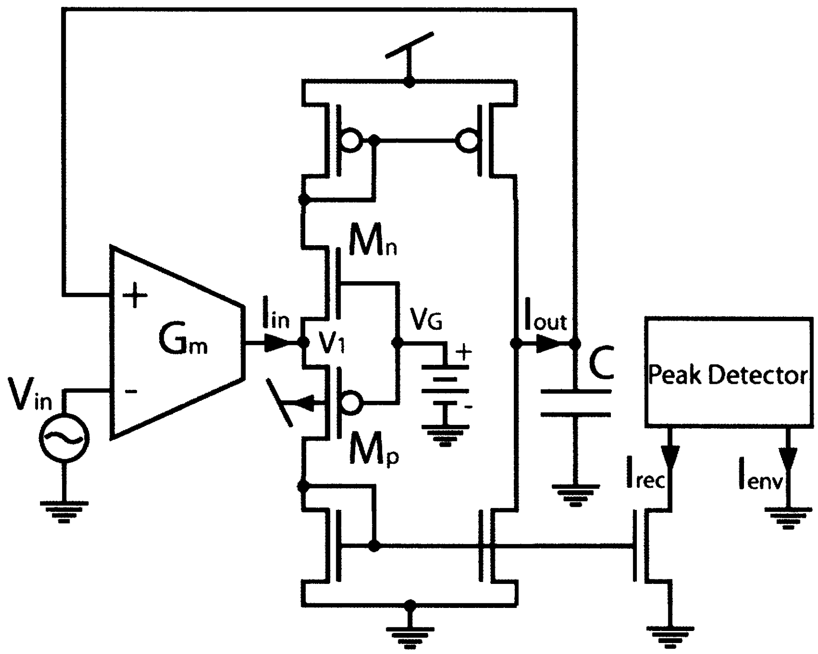


Figure 2: Basic Rectifier Topology.

$I_{out} = -I_{in}$, the voltage across the capacitor is the low-pass filter transfer function: $V_{OUT} = \dfrac{V_{in}}{1 + s\dfrac{C}{G_m}}$.

Then, the current through the capacitor is: $I_{out} = \dfrac{sC \cdot V_{in}}{1 + s\dfrac{C}{G_m}}$. If the pole $\dfrac{G_m}{C}$ is chosen to be sufficiently below the lowest frequency of interest f_{min}=100Hz, we have $I_{out} = G_m \cdot V_{in,AC}$ independent of the input DC voltage or carrier frequency. In this implementation, the rectifier output current I_{rec} is the negative half-wave corresponding to $I_{out} = -I_{in} = G_m \cdot V_{in,AC}$ with ideally zero DC offset. As we have seen, however, there is one very important condition: $I_{out} = -I_{in}$. We will show that both the minimum detectable signal and an observed residual DC offset component of the I_{rec} current are determined by this condition. We have described a different variation of this topology with significantly lower dynamic range in [9].

When designing the rectifier, we would like G_m to be constant over a wide range of input voltages. We also want to avoid tiny input signals that are prone to noise and other effects [10]. These conditions require using wide-linear-range transconductor techniques to implement the G_m

transconductor in Figure 2. These techniques are described in detail in [10]. The topology of the transconductor used in our design is shown in Figure 3 and is hereafter referred to as the WLR OTA. Much of the increase in the input voltage swing of the transconductor comes from using the well rather than the gate as an input in the differential-pair devices. The gates of these devices are connected to their respective drains to implement gate degeneration [10], which further increases the input voltage swing. Transistors B1 and B2 implement bump-linearization techniques [10]. The combination of these techniques allowed us to obtain 1.7Vpp of the input voltage swing. We implement a geometric scaling factor of $N=5$ in the output current mirror arms of the WLR OTA of Figure 3. This scaling improves power consumption, at the cost of worsening noise performance a little, as we discuss later.

2.2.1. Rectifying Class-B Mirror Topology

The implementation of a basic class-B mirror is shown in Figure 2. This structure is capable of sourcing and sinking current from the input I_{in} and mirroring it to the output I_{out}, and is an example of a class of current conveyor circuits. If no current is applied to the input node, the input devices, Mn and Mp, are both turned off. Since the magnitude of the gate-to-source voltages for Mn and Mp must be sufficient to obtain a source or sink current equal to the input current, large voltage swings are required at the input node V_1 to turn these diode-like devices on. Thus, a voltage dead-zone is present at the input node such that no current is mirrored until the node voltage has changed significantly. The deadzone is about 2.2Vpp in the MOSIS 1.5um process, and is comprised of the sum of the NMOS and PMOS diode drops. This dead-zone is typically not a problem for high-current systems that are able to recharge any parasitic capacitance quickly. However, for micropower systems this dead-zone presents a power-speed tradeoff, causing the rectification to fail if I_{in} is unable to recharge the parasitic node capacitance fast enough to turn the input devices Mn and Mp on during some portion of the input cycle. The magnitude of the deadzone is a weak logarithmic function of the input current level, but, for simplicity, we shall assume that it is almost constant.

Dead-zone reduction techniques for class-B mirrors have received attention for signal-processing applications in the recent past [11]. Class AB biasing techniques with output offset-correction to subtract the quiescent bias current have been proposed. We chose to alleviate the dead-zone problem with a combination of an amplification and a class AB biasing technique as shown below.

Assume that $I_{in} = I_0 \cdot \sin(\omega t)$ and that the dead-zone width is a constant V_D peak-to-peak. The parasitic capacitance C_L at the node V_1 consists of two parts: The capacitance C_{node}, due to the output WLR parasitics and node capacitance, and C_p, the gate-to-source parasitics of Mn and Mp. So,

$$C_L = C_{node} + C_p \text{, where } C_{node} \gg C_p \tag{1}$$

Now if the amplitude I_0 is small enough as to be guaranteed not to turn Mn and Mp on, we have

$$V_1 = \frac{I_0}{G_{out} + sC_L} \approx \frac{I_0}{sC_L} \tag{2}$$

where G_{out}, the output conductance of WLR OTA, is very small and may be neglected in the frequency range of interest. The voltage V_1 amplitude increases as we increase I_0. Finally, as the V_1 amplitude approaches $\frac{V_D}{2}$, the rectifier starts to output current. Thus, the minimum detectable I_{in} current is

$$I_{in,MIN} = \omega \cdot C_L \cdot \frac{V_D}{2} \tag{3}$$

Since the maximum possible I_{in} current is the effective bias current of the WLR OTA, $N \cdot I_B$, we obtain a dead-zone output dynamic range limitation in currents D_0 given by the ratio of NI_B to $I_{in,\,MIN}$ to be,

$$D_0 \leq \frac{N \cdot I_B}{\pi \cdot f_{MAX} \cdot C_L \cdot V_D} \tag{4}$$

Since the transconductor is just linear over this range of operation of currents, the dynamic range in input voltages is the same as the dynamic range in the output currents and also given by Equation (4). We notice that we need to spend power by increasing I_B if we desire to have a large dynamic range D_0 or a large frequency of operation f_{max}. In other words, as is commonly observed, power is necessary to get both speed and precision. Equation (4) quantifies our earlier power-speed tradeoff discussion.

Figure 4 illustrates a circuit modification of a basic class-B mirror topology to improve the dead-zone limited dynamic range D_0. Here, the feedback amplifier, *A*, drives the gates of Mn and Mp, thus reducing the voltage swing needed at the V_1 node, and keeping it almost clamped. Again, assuming that I_0 is small enough to not turn Mn and Mp on, we get,

$$V_1 = \frac{I_0}{G_{out} + s(C_{node} + A \cdot C_p)} \approx \frac{I_0}{s(C_{node} + A \cdot C_p)} \tag{5}$$

where $A \cdot C_p$ represents the Miller multiplication of source-to-gate capacitances of Mn and Mp. Now,

$$V_G = A \cdot V_1 \approx \frac{I_0}{s(C_p + C_{node} / A)} \tag{6}$$

and increases as we increase I_0. Finally, as V_G approaches $\frac{V_D}{2}$, the current starts to come out. Thus, the minimum detectable I_{in} current is now given by

$$I_{in,MIN} = \omega \cdot (C_p + \frac{C_{node}}{A}) \cdot \frac{V_D}{2} \approx \omega \cdot C_p \cdot \frac{V_D}{2} \tag{7}$$

provided that the gain *A* is high enough. Now the dead-zone dynamic range limitation is given by

$$D_0 \leq \frac{N \cdot I_B}{\pi \cdot f_{MAX} \cdot C_p \cdot V_D} \tag{8}$$

and constitutes an improvement by a factor of $\frac{C_L}{C_p} = 1 + \frac{C_{node}}{C_p} \gg 1$ over the basic class-B mirror topology. We see that it is important to have the gate-to-source capacitances that constitute C_p be as small as possible to get a large improvement in dynamic range. That's why we use minimum size devices for Mn and Mp, connect the well of the Mp device to V_{DD} rather than to its source although this increases the dead-zone V_D, and operate in subthreshold as far as possible since the only contributor to the gate-to-source capacitances in subthreshold are overlap capacitances in Mn and Mp. Tying the well of the Mp device to V_{DD} increases V_D somewhat, but the decrease in C_p due to the exclusion of C_{gb} is a far more substantial effect, especially on the low end of the dynamic range that we are interested in, where Mn and Mp are in subthreshold, and C_{gb} is the major contributor to C_p.

A further improvement in D_0 is possible by reducing the dead-zone V_D. Figure 5 shows that this reduction can be accomplished by introducing a constant DC voltage shift V_0 between the gates of the Mn and Mp rectifying devices. In this circuit if the I_{in} current is positive, Mp has to be on, so its gate voltage $V_{out,BOT}$ is low enough. The device Mn's gate voltage $V_{out,TOP}$ is higher by V_0, and needs to go up by only $V_D - V_0$ to open Mn as the I_{in} current's sign changes. Therefore, the dead-zone is reduced to $V_D - V_0$. This dead-zone reduction technique is limited because of an upper bound on V_0. From applying the translinear principle, it follows that this technique will result in an output offset current – even with no I_{in} current present, $V_{out,BOT}$ and $V_{out,TOP}$ gate voltages will be set by the A amplifier such that the Mn and Mp standby currents (zero-input currents) are equal. These standby currents have an exponential dependence on V_0 and are mirrored directly to the output of the rectifier stage. We require this zero-input offset current to be no more than a few pA, thus setting a ceiling on V_0 of approximately 1.55V in the MOSIS 1.5um process for minimum size Mn and Mp. It is possible to have dummy devices and subtract some of these standby currents, but as we will discuss later, having a large V_0 where such subtraction would be beneficial is undesirable because of thermal noise rectification. The class AB V_0 technique yields a dead-zone reduction from 2.2Vpp to 0.65Vpp – an improvement of a factor of 3, or 10dB in D_0. Figure 6 illustrates one possible implementation of an A amplifier with the "floating battery" V_0. The value of V_0 can be adjusted to some degree by changing the bias current I_{b2} of the A amplifier.

2.2.2. Theoretical Analysis of Thermal Noise Rectification

We now examine another limitation on the system dynamic range due to the noise of the WLR OTA. For our device sizes and currents the effect of 1/f noise in our circuit is negligible in subthreshold operation [10]. However, the thermal noise current at the WLR OTA output is fed to the class-B mirror, rectified by it, and mirrored to the output, creating a residual output current floor that degrades the minimum detectable signal and dynamic range of the system. The current power spectral density of the white noise at the WLR OTA output is

$$\bar{i}_{noise}^{2}(f) = n \cdot q \cdot NI_B \tag{9}$$

where,

$$n = \left(\frac{2 \cdot \kappa_n}{\kappa_p + \kappa_n} \right)^2 \cdot N + 2N + 2 \approx 2.68N + 2 \tag{10}$$

represents the effective number of noise sources in our WLR OTA, κ_n is the subthreshold exponential parameter of the NMOS transistors in the current mirror of Figure 3, and κ_p is the subthreshold exponential parameter of the differential-pair PMOS transistors. Details of how to compute the effective number of noise sources in such circuits are provided in [10].

From our previous discussion about the dead-zone limitation, it is clear that the higher the frequency of the input current, the higher the threshold presented by the dead-zone

$$I_{in,MIN} = \pi \cdot f \cdot C_p \cdot (V_D - V_0) \tag{11}$$

Therefore, almost all of the low-frequency part of the white noise spectrum passes to the output, whereas the high-frequency part gets filtered out by the capacitor C_p. For simplicity, we shall assume that the dead-zone and C_p create a low-pass filter with an infinitely steep slope at a still-to-be-determined cut-off frequency f_0. With this assumption, the class-B mirror behaves as if the I_{in} current were Gaussian with zero mean and

$$\sigma^2 = n \cdot q \cdot NI_B \cdot f_0 \tag{12}$$

Then,

$$\begin{aligned} \bar{I}_{rec} &= \int_0^{+\infty} I \cdot \frac{1}{\sqrt{2\pi} \cdot \sigma} e^{-\frac{I^2}{2\sigma^2}} \cdot dI \\ &= \frac{\sigma}{\sqrt{2\pi}} = \sqrt{\frac{n \cdot q \cdot NI_B \cdot f_0}{2\pi}} \end{aligned} \tag{13}$$

To estimate the cut-off frequency f_0 we note that once the frequency-dependent threshold presented by the dead-zone in Equation (11) gets higher than the σ of Equation (12), little current is output by the rectifier. Therefore, a reasonable estimate is to assume that the frequency-dependent threshold at f_0 is at σ. Thus,

$$\pi \cdot f_0 \cdot C_p \cdot (V_D - V_0) \cong \sqrt{n \cdot q \cdot NI_B \cdot f_0}$$
$$\Rightarrow \sqrt{f_0} \cong \frac{\sqrt{n \cdot q \cdot NI_B}}{\pi \cdot C_p \cdot (V_D - V_0)} \tag{14}$$

Plugging the result for f_0 back into Equation (13), we obtain

$$\bar{I}_{rec} \cong \frac{n \cdot q \cdot NI_B}{\pi\sqrt{2\pi} \cdot C_p \cdot (V_D - V_0)} \tag{15}$$

Recalling Equation (8) for the dead-zone dynamic range limitation, we have

$$\bar{I}_{rec} \cong \frac{n \cdot q \cdot f_{MAX} \cdot D_0}{\sqrt{2\pi}} \tag{16}$$

In our design, $N = 5 \Rightarrow n \approx 15.4$, $q = 1.6 \cdot 10^{-19} C$, D_0 was designed and simulated to be $80dB = 10^4$ for $f_{MAX} = 10kHz$, $I_B = 200nA$ (bias current through WLR OTA), and $I_{b2} = 200nA$ (bias current through A amplifier yielding V_0 = 1.55V and a deadzone of 0.65V pp). That gives us $\bar{I}_{rec} \cong 100pA$. The corresponding experimentally measured result, which we present in Section IV, is $\bar{I}_{rec} = 119pA$, indicating that our approximations and assumptions are sound.

The implication of Equation (16) is that the larger we make D_0 to increase the minimum detectable signal limited by the dead-zone non-linearity, the higher the rectified-noise-current floor becomes, and the greater is the degradation in minimum detectable signal caused by this current floor. Since the overall dynamic range of the system is determined by whichever effect yields a larger minimum detectable signal (dead-zone limitation or noise-rectification), the maximum dynamic range is achieved if both effects yield the same limit. At this optimum, we are spending as much power as necessary to achieve the highest D_0 possible but not so much power that the rectification-noise-floor increases and limits the dynamic range to values below D_0. Alternatively, at a fixed power level, if the deadzone and noise-rectification limits match, the deadzone is at a small enough value such that we can overcome it with faint amplified signals but not so small that the rectified-noise-current floor swamps the output current due to the faint signals. Thus, the optimum dynamic range is achieved when the limit of minimum detectable signal due to the rectified-noise-current floor of Equation (15) becomes equal to the mean value of the dead-zone minimum detectable current. The dead-zone minimum detectable current is a half-wave rectified sinusoid with an amplitude given by Equation (11). If we realize that a half-wave-rectified sine wave has a mean current that is 1/□ of its amplitude, and use Equations (15) and (11) we find at the optimum that

$$\frac{n \cdot q \cdot NI_B}{\pi\sqrt{2\pi} \cdot C_p \cdot (V_D - V_0)} = \frac{\pi \cdot f_{MAX} \cdot C_p \cdot (V_D - V_0)}{\pi} \tag{17}$$

Algebraic simplification yields

$$\left[C_p \cdot (V_D - V_0)\right]_{optimum} = \sqrt{\frac{n \cdot q \cdot NI_B}{\pi\sqrt{2\pi} \cdot f_{MAX}}} \tag{18}$$

Substituting this result back into Equation (8), we obtain

$$D_{optimum} = \sqrt[4]{\frac{2}{\pi}} \cdot \sqrt{\frac{NI_B}{n \cdot q \cdot f_{MAX}}} \tag{19}$$

From Equation (18) we see that the optimal dynamic range depends only on topological parameters like n and N, the charge on the electron q, and is independent of technological parameters like C_p and V_D. To get more dynamic range at a given f_{max} and in a given technology, we must spend more power according to Equation (19), and simultaneously decrease V_0 in Equation (18) to ensure that we are at the optimum. Intuitively, we burn power to allow smaller and smaller signals to break the deadzone but concomitanly increase the deadzone such that the noise-rectification limit always matches the deadzone limit.

In our design, due to the power constraints, we can only afford $I_B = 200nA$. According to Equation (19), that gives us a maximum possible system dynamic range of $D_{optimum} \approx 75dB$. In order to reach this optimum we decrease V_0, and increase the deadzone, by turning down the bias current I_{b2} of the A amplifier. In section 2.4, we show experimentally, that we can actually achieve this theoretical optimum.

2.3. Peak Detector Design

Figure 7 illustrates a simple current peak detector topology described in [4]. We will just highlight some nuances of its operation since they are important to the discussion of a better peak detector presented in this paper. As I_{in} increases (the "attack"), it discharges parasitic capacitance C_{par} decreasing the V_1 voltage. The decrease in V_1 causes transistor M1 to open and to quickly decrease the V_2 voltage almost instantaneously (we have a very fast attack time constant). The decrease in V_2 increases I_{out} and also increases the drain current of M2 to a point where it equals I_{in} via negative-feedback action. The phase margin of this feedback loop determines overshoot of the output current I_{out} during the attack. As I_{in} decreases (the "release"), drain current in M2 quickly increases the V_1 voltage across parasitic capacitance C_{par}. The increase in V_1 causes transistor M1 to turn off. Now V_2 changes due to the charging of C_r by I_r. This change is linear, i.e., $C_r \frac{dV_2}{dt} = I_r \Rightarrow V_2 = V_{2,0} + \frac{I_r}{C_r} t$. The dynamics of V_2 yields an expression for decay of the output current (M3 drain current expression for the weak inversion) during release $I_{out} \propto e^{-\frac{\kappa V_2}{\phi_t}} \propto e^{-\frac{\kappa I_r}{\phi_t C_r} t}$.

Since the definition of the release time constant is obtained from $I_{out} \propto e^{-t/\tau_r}$, we have

$$\tau_r = \frac{C_r \cdot \phi_t}{\kappa \cdot I_r} \tag{20}$$

where $\phi_t = \frac{k \cdot T}{q} \approx 25mV$, and κ is the subthreshold exponential parameter of the PMOS transistors. We now analyze the feedback loop inherent in Figure 7. The block diagram of this feedback loop is shown in Figure 8 and is based on standard small-signal parameters of the transistors M1 and M2. Taking Equation (20) into account, the loop transmission is given by

$$L(s) = -\frac{1}{1 + s \cdot \tau_r} \cdot \frac{A_2}{1 + s \frac{C_{par}}{g_{ds2}}} \tag{21}$$

where $A_2 \equiv \frac{g_{m2}}{g_{ds2}}$. We have ignored capacitances between nodes V_1 and V_2 in our analysis. The Bode plot of the loop transmission is shown on Figure 9. The criterion for good phase margin in the feedback loop (45 degrees or more) is that $\frac{A_2}{\tau_r} < \frac{g_{ds2}}{C_{par}}$, which can be rewritten as

$$I_{in} > I_r \cdot \frac{C_{par}}{C_r} \cdot A_2^2 \tag{22}$$

We see that the dynamic range of good-phase-margin operation of this peak detector is limited to large currents even for modest values of A_2.

Figure 10 illustrates a standard current-mode low-pass filter topology. For a review of the ideas behind current-mode filtering, see [12]. Again, we will highlight nuances of its operation crucial to the discussion of our peak detector functioning. The time constant of this filter is [12]

$$\tau_a = \frac{C_a \cdot \phi_t}{\kappa \cdot I_a} \tag{23}$$

Transistor M1 converts the input current into its logarithm. Transistor M2 performs dynamic translinear low-pass filtering, such that its source voltage is proportional to the logarithm of the low-pass filtered input current. This voltage is then shifted by M3 to keep the gain of this structure close to unity, and then expanded by M4 to convert a logarithmic voltage into an output current.

Figure 11 shows our novel current-mode peak detector topology with wide-dynamic-range nanopower operation. The attack and release time constants are adjustable. First, we note that the current through the M3 transistor is always equal to I_a, provided that the parasitic capacitance of the V_1 node is small. Therefore, like in current mode low-pass filter, the source voltage of M2, V_0, is proportional to a logarithm of the low-pass-filtered input current with a time constant given by Equation (23). The M3 transistor, however, only acts like a simple shifter during attack: As I_{in} increases during an attack phase, the V_0 voltage decreases. This decrease causes the drain current of M3 to decrease. The I_a current then quickly discharges parasitic capacitance C_{par} decreasing V_1. The decrease in V_1 causes transistor M5 to open and to quickly decrease V_2, thus restoring M3's drain current. Since M3 does behave like a shifter during attack, the attack time constant is given by Equation (23). The feedback loop formed by M5 and M3 is similar to the one in the simple peak-detector topology of Figure 7, and has already been analyzed. To provide good phase margin, the current I_a still has to satisfy Equation (22), but now the good-phase-margin conditions do not affect the dynamic range of operation, because all currents in the M3-M5 feedback loop are fixed. Thus, we may pick current values in the loop to give us good phase margin for all inputs. As I_{in} decreases during release, the V_0 voltage goes up. This causes the drain current of M3 to increase, increasing the V_1 voltage, which turns off transistor M5. Now, V_2 only changes due to charging of C_r by I_r such that the release time constant is given by Equation (20).

The peak-detector topology of Figure 11 does experience a slight dependence of its output current on frequency: The ripple at the V_0 node after attack filtering is larger for low carrier frequencies than for high frequencies. Consequently, the following release filter will follow the peaks of the ripple around the frequency-independent V_0 mean, and cause a slight rise in the output current for low frequencies.

2.4. Experimental Results

A chip with this envelope detector was fabricated on AMI's 1.5um CMOS process through MOSIS. Figure 22 shows a photograph of the die.

Figure 12 shows experimental waveforms of the rectifier output current at $f = 100Hz$ for a tone-burst input. The half-wave rectification is clearly evident. Figure 13 shows experimental waveforms of the envelope-detector output current for three tone-burst carrier frequencies of 300Hz, 1kHz, and 10kHz with the same input signal amplitude. We can see that the attack time constant is approximately 10ms, and the release time constant is approximately 100ms. Both these time constants may be adjusted by altering I_a and I_r in Figure 11. We do observe more ripple for low-frequency inputs than high-frequency inputs and a weak dependence of the output current as well.

Figure 14 shows experimentally measured envelope detector characteristics at 100Hz, 1kHz, and 10kHz for input signal amplitudes ranging over the entire 75dB of operation. The plot saturates at $V_{in} \approx 1.7Vpp$ on the high end of the dynamic range, and flattens out at approximately $V_{in} \approx 300\mu Vpp$ on the low end, revealing that the envelope detector provides proportional and linear information about the input signal envelope over a dynamic range of 75dB at all audio frequencies of interest. The saturation is caused by the WLR OTA moving out of its linear range while the flattening is due to the thermal-noise-rectified output current floor that we discussed in Section II.

Figure 15 shows experimentally measured envelope detector characteristics at 10kHz for various I_{b2}, i.e., various dead-zone widths. At low values of I_{b2}, the dead-zone is wide, implying that both the dead-zone-limited dynamic range and the rectified-noise current floor are low. By increasing I_{b2} we may decrease the dead-zone width, improving the dead-zone-limited dynamic range, but also increasing the rectified-noise current floor. At the optimal point ($I_{b2} = 25nA$) the dead-zone minimum detectable signal equals the rectified-noise current floor, and we obtain 75dB of dynamic range, in excellent agreement with the theory of Section II. Further increases in I_{b2}, i.e., reductions in dead-zone width, lead to improvement of the dead-zone minimum detectable signal, but degrade the rectified-noise current floor, degrading overall dynamic range of the system. Figure 16 illustrates this point further, showing the overall dynamic range of the system vs. I_{b2}.

Figure 17 shows the rectified-noise current floor measured at the output of the class-B NMOS mirror ($I_{res,NMOS}$), PMOS mirror ($I_{res,PMOS}$), and the output of the peak detector ($I_{res,PD}$), as the WLR OTA bias current, I_B, varies. As we would expect, all three currents are almost identical. The data also reveal that the peak detector contributes little to the noise of the whole system.

Figure 18 confirms the independence of the rectified-noise current floor from the input DC voltage over a wide range of operation. This result is consistent with the theory of Section II and also reveals the insensitivity to the input DC voltage of our system. The output current of the system was also invariant with the input DC voltage but we have not shown this data.

Figure 19 shows the output current floor measured for various I_{b2} A-amplifier biases, i.e. for various dead-zone widths. Although it was impossible to measure the dead-zone width quantitatively, we observed qualitative agreement between this experimental result and Equation (15).

Figure 20 confirms that the rectified-noise current floor is invariant across several fabricated chips and not a parasitic "leakage" effect but a fundamental one due to thermal noise. We see that the slope of the lines is different from unity, implying that the output noise floor has a slightly nonlinear dependence on I_B instead of the purely linear dependence predicted by Equation (15). This nonlinearity may be explained by the lowering of the number of effective noise sources, i.e. n in Equation (15) as the increasing WLR OTA bias current I_B causes a transition from subthreshold operation into moderate-inversion or strong-inversion operation. Such effects have also been described in the measurements described in [10].

Finally, we performed an experiment to estimate n experimentally: We lowered the WLR OTA bias current I_B significantly, effectively lowering its own thermal noise to small levels. Then, we input a white-noise voltage into the envelope detector and measured the output current. The input now creates the rectified-noise current floor rather than the internal white noise. Figure 21 shows the output current vs. generator voltage $\sqrt{\overline{\upsilon}^2}$ for $I_B = 20nA$ and $I_B = 40nA$. We observe a leveling off of the output current floor at low input voltages due to the intrinsic internal white noise of the WLR OTA. We can "map" I_B from Figure 20 to $\sqrt{\overline{\upsilon}^2}$ from Figure 21 such that they produce exactly the same output current "noise floor". This mapping means that the current spectral power on the output of the WLR OTA would have to be the same in both cases, i.e.

$$
\begin{aligned}
& n \cdot q \cdot N I_{B,20} = n \cdot q \cdot N I_{B,21} + \overline{\upsilon}^2 \cdot \left(\frac{N \cdot I_{B,21}}{V_L} \right)^2 \\
& \Rightarrow n = \frac{\overline{\upsilon}^2 \cdot \left(\frac{N \cdot I_{B,21}}{V_L} \right)^2}{q \cdot N \cdot (I_{B,20} - I_{B,21})}
\end{aligned}
\tag{24}
$$

where $N = 5$, $V_L = 0.85V$, $I_{B,21} = 20nA$ (we used the right curve in Figure 21).

From the experiment we estimate that $n \approx 25$, which is in reasonable agreement with our theoretical calculations of $n = 15.4$

2.5. Conclusions

The combination of a wide-linear-range transconductor topology, a modified class-B current mirror, and a novel current-mode peak-detector yielded a 75dB 2.□□W envelope detector with frequency-independent operation over the entire audio range from 100Hz to 10kHz. The current-mode peak detector provided wide-dynamic-range good-phase-margin operation with adjustable attack and release time constants. We confirmed theoretical predictions of the minimum detectable signal of the envelope detector due to dead-zone-limiting effects and thermal-noise-rectification effects experimentally. We also achieved the maximum possible dynamic range predicted from theory. The detector should be useful in ultra low power bionic implants for the deaf, hearing aids, and low-power speech-recognition front ends where automatic gain control and spectral-energy computations require the use of envelope detection. The topology of the detector could also potentially be useful in higher-frequency applications like sonar or RF-demodulation if more power is consumed.

2.6. 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] W. A. Serdijn, A. C. van der Woerd, J. Davidse, H. M. van Roemund, "A Low-Voltage Low-Power Fully-Integratable Front-End for Hearing Instruments," *IEEE Transactions on Circuits and Signals – I: Fundamental Theory and Applications*. Vol. 42, No. 11, November 1995.

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

[4] 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.

[5] 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

[6] 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.

[7] T. Lu, M. Baker, C. Salthouse, J. J Sit, S. Zhak, and R. Sarpeshkar, "A Micropower Analog VLSI Processing Channel for Bionic Ears and Speech Recognition Front Ends", *IEEE Intl. Symposium on Circuits and Systems,* 2003.

[8] B. Gilbert, "Translinear Circuits: A Proposed Classification," *Electronics Letters,* vol. 11, no. 1, pp. 14–16, 1975.

[9] M. Baker, S. Zhak, and R. Sarpeshkar, "A micropower envelope detector for audio applications," *IEEE Intl. Symposium on Circuits and Systems,* 2003.

[10] 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.

[11] T. Kurashina, S. Ogawa, K. Watanabe, "A High Performance Class AB Current Conveyor," *IEEE International Conference on Electronics Circuits and Systems 1998,* Vol. 3, pp. 143-146.

[12] D. R. Frey, "Log-domain filtering: an approach to current-mode filtering", *IEE Proceedings, Part G*, vol. 140, pp. 406-416, 1993.

Error! Objects cannot be created from editing field codes.

Figure 3: A Wide-Linear-Range Transconductor (WLR OTA) [10].

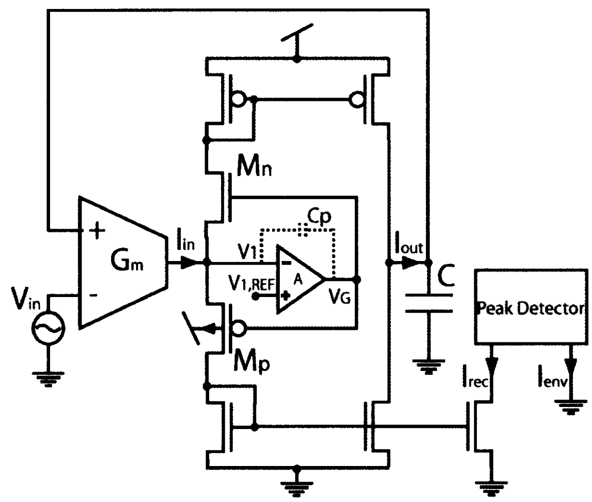


Figure 4: Class-B Current Mirror with Active Feedback.

Error! Objects cannot be created from editing field codes.

Figure 5: Modified Class-B Mirror with Active Feedback and Dead-Zone Reduction.

Error! Objects cannot be created from editing field codes.

Figure 6: Active Feedback Amplifier with "Floating Battery" Implementation.

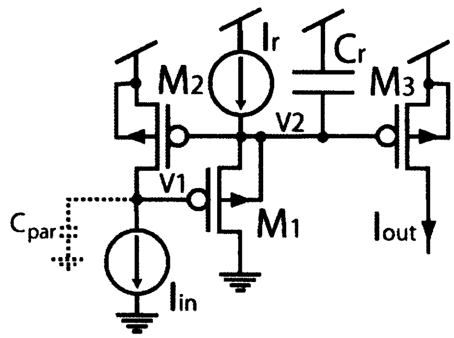


Figure 7: Simple Current Peak Detector [4].

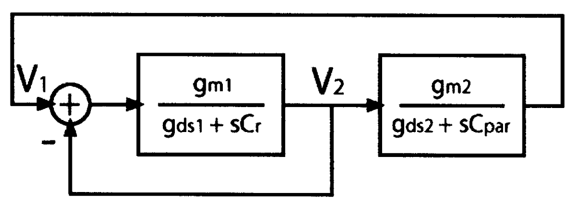


Figure 8: Block Diagram of the Small-Signal Feedback Loop of Figure 7.

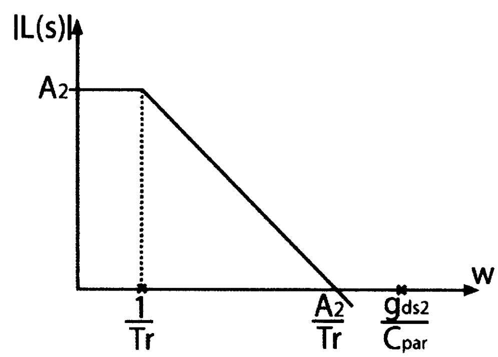


Figure 9: Bode Plot of the Loop Transmission of Figure 8.

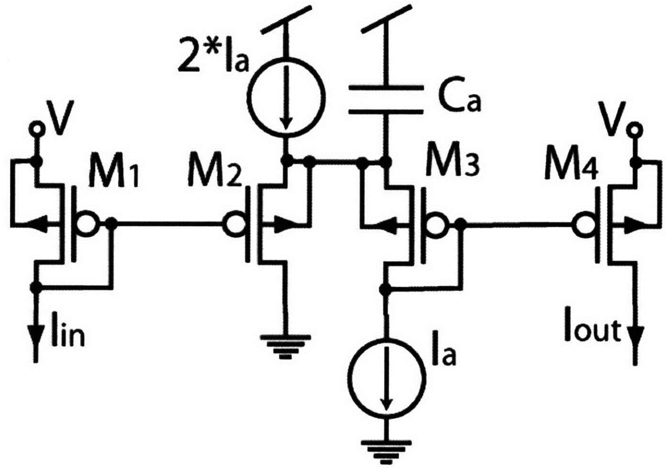


Figure 10: Current Mode Low-Pass Filter [12].

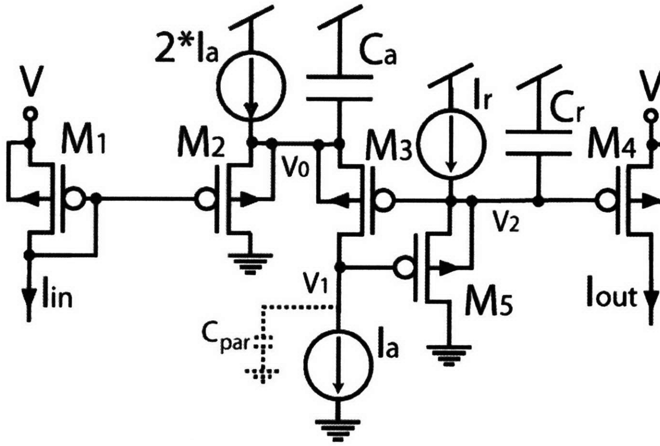


Figure 11: Wide Dynamic Range Current-Mode Peak Detector with Adjustable Attack and Release Time Constants.

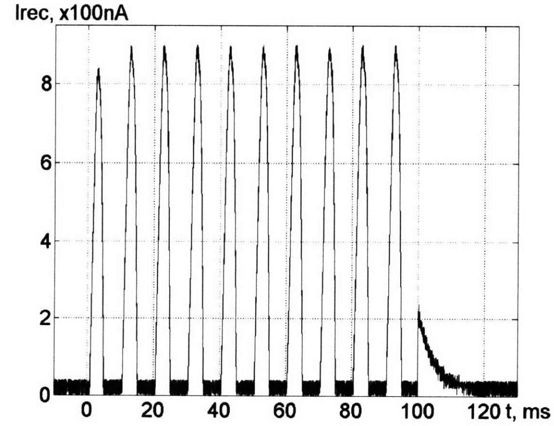


Figure 12: Experimental Rectifier Output Current Waveform.

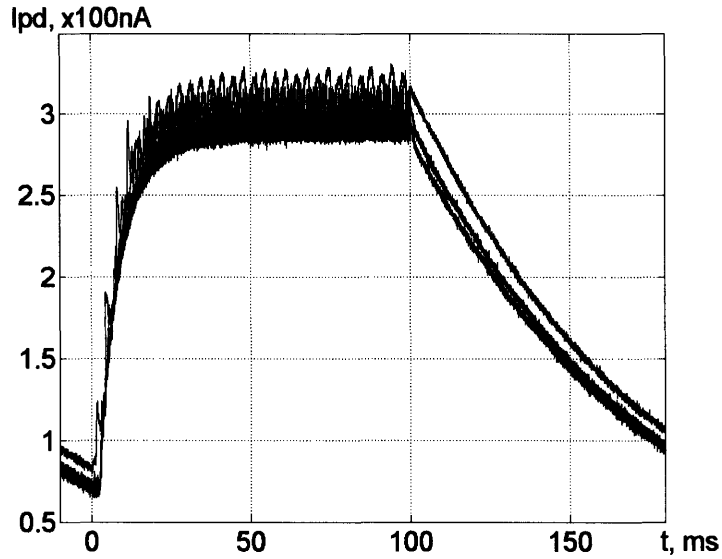


Figure 13: Experimental Peak Detector Output Current Waveform;

f=100Hz, 1kHz, and 10kHz; Ta=10ms, Tr=100ms.

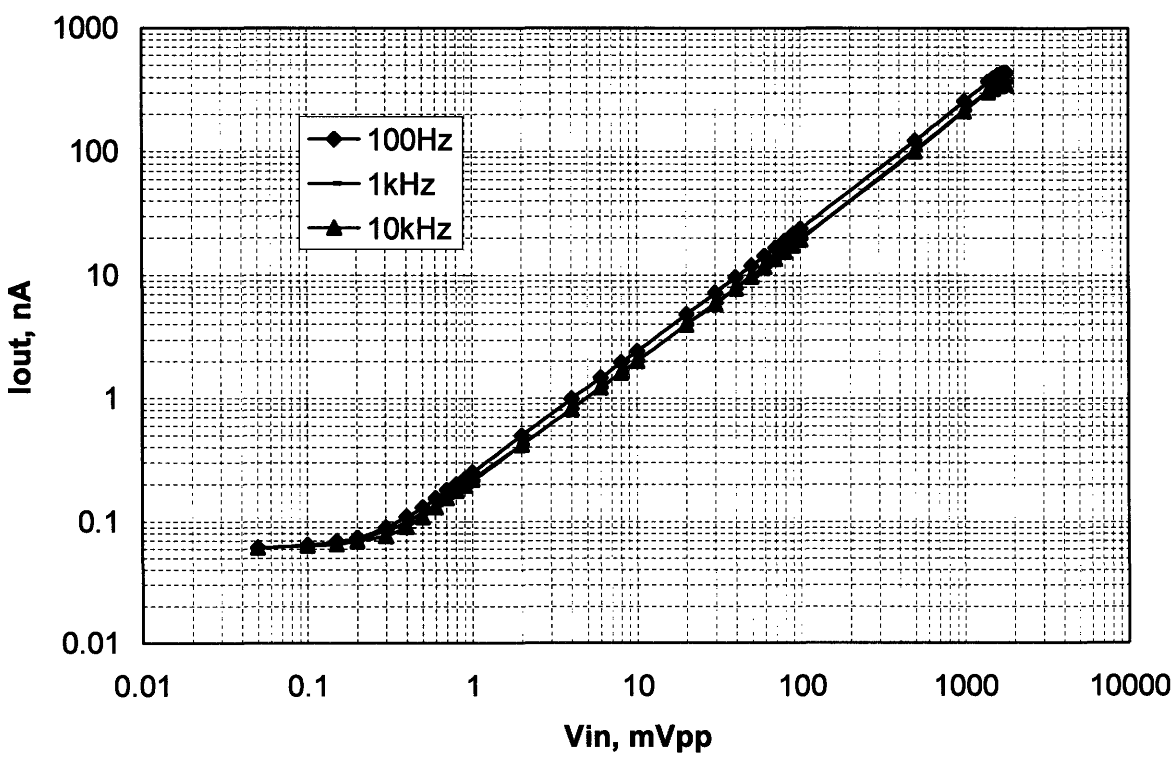


Figure 14: Experimentally Measured Envelope Detector Characteristics;
f=100Hz, 1kHz, and 10kHz.

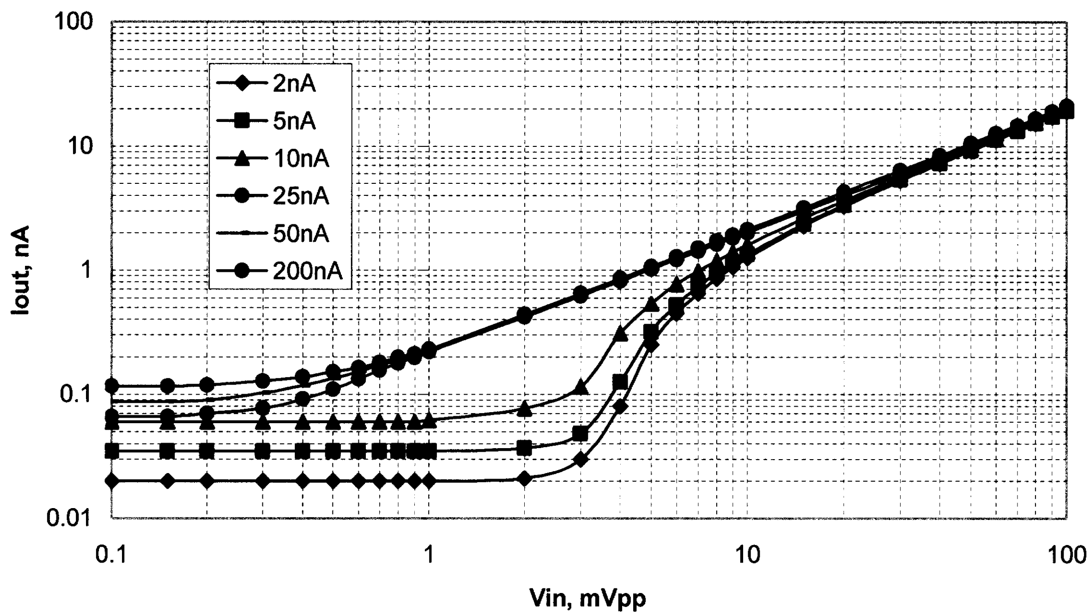


Figure 15: Experimental Envelope Detector Characteristics for f=10kHz, and Various Ib2, i.e. Dead-Zone Widths.

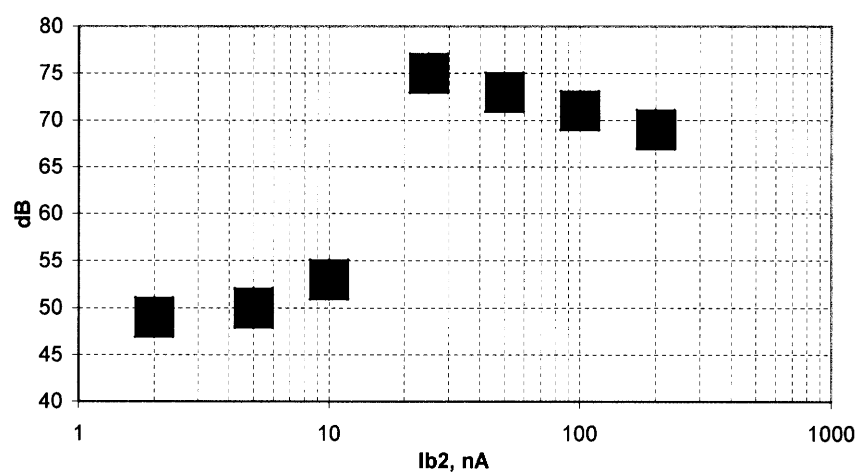


Figure 16: Overall Dynamic Range for Various Ib2, i.e. Dead-Zone Widths.

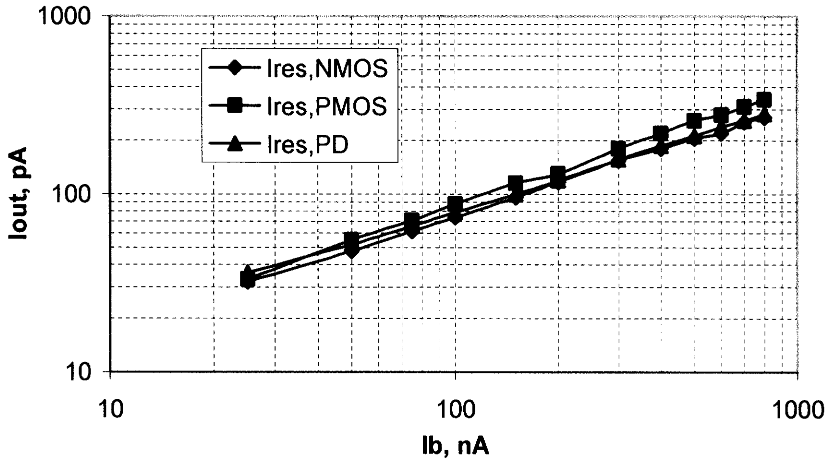


Figure 17: Output Current vs. WLR OTA Bias Measured at NMOS and PMOS Rectifier Outputs, and PD Output.

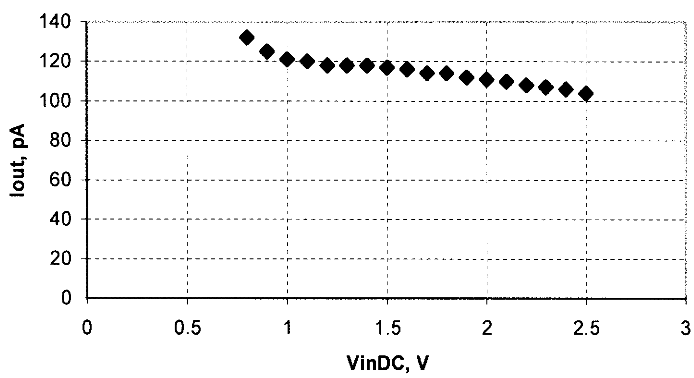


Figure 18: Output Current vs. Input DC Voltage.

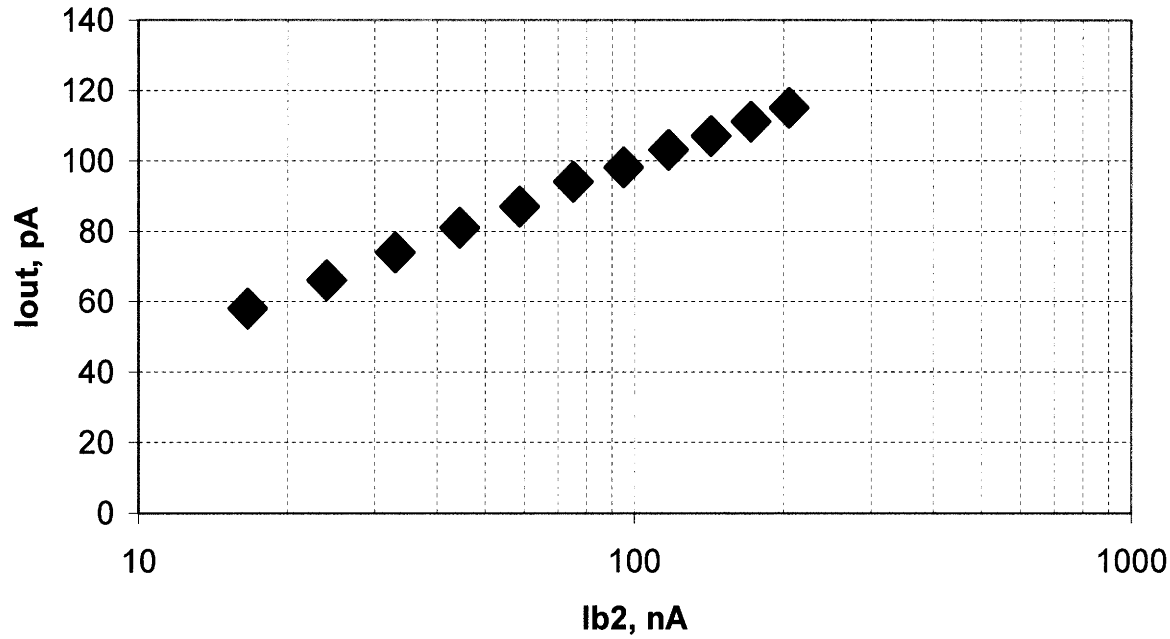


Figure 19: Output Current for Various Ib2, i.e. Dead-Zone Widths.

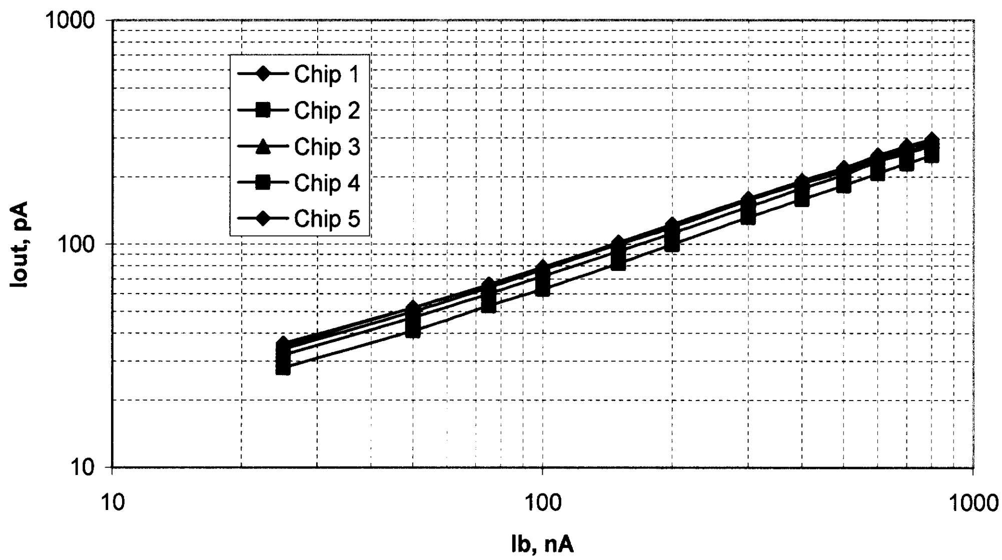


Figure 20: Output Current vs. WLR OTA Bias Across Several Fabricated Chips.

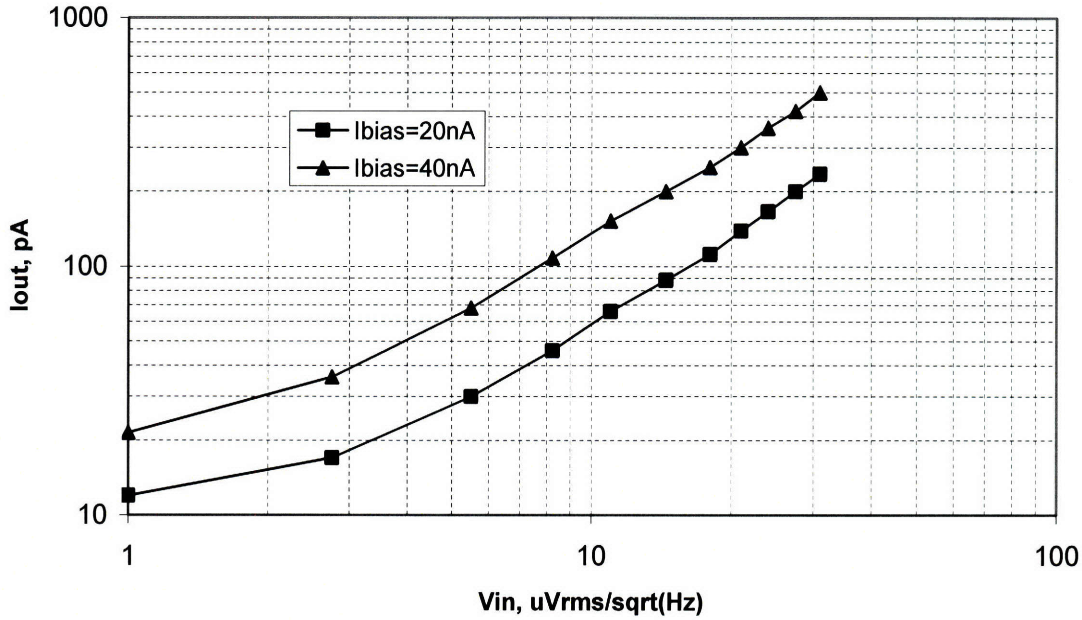


Figure 21: Output Current vs. White Noise Generator Voltage.

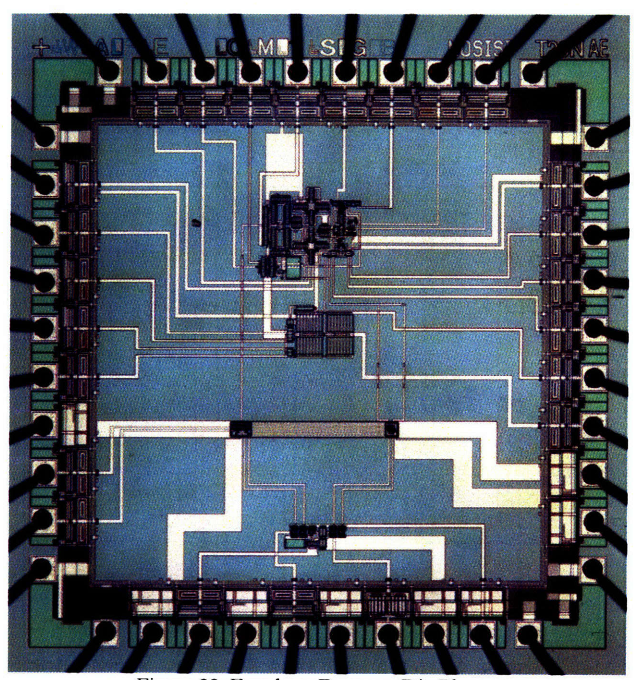


Figure 22: Envelope Detector Die Photo.

3. Single-mode one-dimensional transmission-line cochlear architectures

Abstract—**Rational approximations to distributed irrational partition impedances in the biological cochlea are shown to capture its essential features: These include good frequency- selectivity and steep roll-offs, high active amplification, peak-frequency shifts and gain adaptation with stimulus level, and invariance of the fine time structure of the impulse response with stimulus intensity. It is shown that the terminating impedance at the end of the cochlea is crucial for its stability, and various analytic methods for termination are discussed. These termination techniques may suggest why the biological cochlea tapers its amplification at the apex. A composite cochlear architecture that is a good approximation to a transmission-line model is derived with cascaded second-order filters. Our results enable efficient cochlear implementations to be designed as front ends for speech recognition, cochlear implants, or RF spectrum analyzers.**

Index Terms—**Cochlear models, rational impedance, stability, cascade implementation.**

3.1. INTRODUCTION

THE biological cochlea has remarkable sensitivity, frequency selectivity, high gain at the peak frequency or the "best place" of the response, steep roll-offs, a broad frequency range of operation over approximately 3 decades (10 octaves), and an input dynamic range that spans 6 orders of magnitude in sound pressure. The amplification mechanism in the cochlea is nonlinear, compressing a wide input dynamic range into a much narrower output dynamic range in the auditory nerve fibers by reducing the sensitivity and selectivity of the cochlear amplifier with stimulus intensity. The peak (or best) frequency of the response exhibits an approximately "half-octave shift" towards lower frequencies at high stimulus intensities [1]. In addition, the fine time structure of the impulse response remains relatively invariant with stimulus intensity [2]. In a healthy ear, the fine time structure of the cochlear response is represented in the temporal discharge patterns of auditory nerve fibers for frequencies up to 4 kHz [3], [4].

Cochlear implementations are useful as front ends for various applications including speech recognition, cochlear implants, and RF spectrum analysis [5]-[7]. The collective amplification and exponential taper of the cochlea provide a very efficient method for implementing a high-resolution spectrum analyzer with a wide frequency range of operation [8]. The nonlinear gain-control and tone-to-tone suppression properties of the cochlea allow good preservation of spectral peaks in the signal, naturally enhancing channel outputs with high signal-to-noise ratio (SNR) with respect to neighboring channels [9]. Speech and other sound stimuli contain information in both their slowly varying envelope and their rapidly varying fine time structure [10]. Frequency modulation derived from the fine time structure can be helpful for speaker identification, music perception [11], and tonal language perception. The precise encoding of stimulus phase in bilateral auditory systems is useful for sound localization and spatial separation of the sound sources, an important cue for localizing signals in noisy environments. Future cochlear implants may attempt to include fine time structure information in the signals delivered to the stimulating electrodes [12], [13]. Therefore, preserving the fine time structure of the speech processor's response with stimulus intensity is needed. Future processors in a cochlear implant or in speech-recognition front ends would benefit from a simple cochlear model that realistically reproduces the features that we have described [14]-[16]. Furthermore, such cochlear models are amenable to very low power analog implementations [6], an important consideration in portable speech processors.

Various cochlear models have been developed. An incomplete list of these might include [17]-[28]. The emphasis in such models has been in understanding the biology rather than in implementation efficiency. One of the earliest attempts to build an active cochlear transmission-line model was to derive theoretically the hydrodynamic impedance of the cochlear partition (CP) based on the experimental data [29]. This model achieves its high amplification and realistic cochlear response by zeroing CP impedance over an appreciable spatial region in the vicinity of the peak (best place) of the response. This model is also one of the simplest – it is one-dimensional and single-mode. However, it contains a pure delay term that can not be implemented with a finite number of lumped elements. Another cochlear model that produces a realistic response is described in [30]. This model assumes that the main effect of the cochlear amplifier is to reduce the CP impedance, which increases CP motion even without an increase in pressure difference across the CP. The model achieves amplification without requiring extra energy to be injected into the traveling wave. It is simple, one-dimensional and single-mode, with a rational form of the CP impedance, although it is not amenable to easy circuit implementations. It also does not exhibit shifts in the peak frequency with input stimulus level, nor does it reproduce the approximate invariance of the fine time structure of the response with stimulus intensity.

The model presented in [29] was extended in [31] with three forms of the intensity dependence of the CP impedance. Intensity-invariance of the fine time structure of each model response was examined by applying the EQ-NL theorem [25], [32] and [33], which replaces the difficult problem of investigating one nonlinear cochlear model with the simpler problem of investigating a family of linear cochlear models.

In section 3.2 we develop a very simple single-mode one-dimensional cochlear model with a rational form of CP impedance, which still captures the important features of the biological cochlear response such as high selectivity and sensitivity at the best frequency, and steep roll-off beyond it. We show that such nonlinear effects as shift in the peak frequency and the near-invariance of the fine time structure of the response with intensity are preserved. In section 3.3 the time-domain stability of this model is investigated. We show that the terminating impedance at the end of the cochlea is crucial for its overall stability, an issue that has not been studied in much detail in prior work but that is extremely important for hardware and software implementations. We discuss various methods for apical termination. Our work on termination might provide insight into the operation of the biological cochlea. In section 3.4 we derive a composite cochlear architecture composed of a cascade of second-order filters from the model of section 3.2. This architecture solves the problem of excessive group delay and excessive compression seen in earlier composite architectures built with second-order filters [6] but essential features of the cochlear response are still preserved. We summarize our contributions in section 3.5.

3.2. Cochlear Model with Rational Shunt Impedance

We adopt the classical point-impedance model of the cochlea. The fluid flow is approximated to be one-dimensional. This approximation is valid if the wavelength of the traveling wave is large compared to the cross-sectional dimensions of the scalae. Therefore, the model can be represented by a transmission line divided into a number of sections. The series inductance represents longitudinal fluid coupling, and the shunt branch represents the CP impedance. The pressure difference P across the CP and the volume velocity U satisfy the transmission-line equations:

$$-\frac{\partial P}{\partial x} = j\omega L(x)\cdot U \tag{1a}$$

$$-\frac{\partial U}{\partial x} = V_{CP} = \frac{P}{Z(j\omega,x)} \tag{1b}$$

where $L(x)$ is the per-length fluid mass in the scalae, $1/Z(j\omega,x)$ is the per-length point-admittance of the CP, and $V_{CP}(j\omega,x)$ is the linear velocity of the CP motion. Equation (1a) describes macromechanical longitudinal fluid coupling in the cochlea, and (1b) represents CP micromechanics. Note that the linear velocity of the CP motion and the volume velocity of the cochlear fluid motion have different dimensions but differ only by a constant factor corresponding to the area of the partition represented by a single lumped section. The definition of the cochlear transfer function (TF) is:

$$TF(j\omega,x) \equiv \frac{V_{CP}}{U(0)} = \frac{1}{U(0)}\frac{P}{Z(j\omega,x)} \tag{2}$$

We assume local scaling symmetry [29], which implies that rather than depending on position and frequency independently, CP impedance $Z(j\omega,x)$, velocity U, pressure difference P and cochlear TF depend only on the following combination of x and ω:

$$\beta(x,\omega) \equiv \frac{\omega}{\omega_c(x)} = \frac{\omega\cdot e^{x/l}}{\omega_c(0)} \qquad s \equiv j\beta \tag{3}$$

where $\omega_c(x)$ is the CF at the location x along the CP, and l is the space constant or characteristic length of the exponential cochlear taper; these parameters define the cochlear position-frequency map. Equations (1a) and (1b) lead to the following ordinary differential equation (ODE) for the pressure difference P:

$$\frac{d^2P}{ds^2} = k^2(s)\cdot P \tag{4a}$$

where

$$k^2(s) \equiv \frac{(4N)^2}{s\cdot Z_n(s)} \tag{4b}$$

and

$$(4N)^2 \equiv \frac{l\cdot L(0)}{L_{CP}(0)/l} \tag{4c}$$

The mass of the cochlear partition at the beginning of the transmission line is $L_{CP}(0)$; the impedance $Z_n(s)$ is dimensionless and obtained by normalizing $Z(j\omega,\ x)$ by $\omega_c(0)L_{CP}(0)$; N is a dimensionless constant approximately equal to the total number of wavelengths of the traveling wave on the CP [29]. Note that $L(0)/L_{CP}(0)$ has units of $1/l^2$ because $L(0)$ is defined in terms of volume velocity while $L_{CP}(0)$ is defined in terms of linear velocity.

The boundary condition that P remains finite as $\beta\rightarrow\infty$ implies that only the forward-traveling wave is present [34]. If we assume that the properties of the cochlea scale slowly relative to the wavelength of the traveling wave, the analytical WKB-approximate solution for the pressure difference P is given by,

$$P(s) \propto k^{-1/2}(s)\cdot \exp\left(-\int_0^s k(s')\,ds'\right) \tag{5}$$

From (2) and (4b), the WKB-approximate solution for the cochlear TF then becomes [29]:

$$TF(s) \propto s\cdot k^{3/2}(s)\cdot \exp\left(-\int_0^s k(s')\,ds'\right) \tag{6}$$

Ignoring the pre-exponential terms, which change slowly with $s=j$□ in comparison with the exponential term, taking the logarithm and then derivative of (6), we get,

$$k(j\beta) \approx -\frac{d}{d\beta}Phase\{TF(\beta)\} + j\cdot\frac{d}{d\beta}\log\left|TF(\beta)\right| \tag{7}$$

From measured cochlear transfer functions of the gain and phase, we can calculate $k(j\beta)$ from (7). It is worth noting that the long-wave approximation and the use of WKB approximation are not necessary for determining $k(s)$ from experimental data: Shera provided a method [35] for obtaining the wave number $k(j\beta)$ in a two-dimensional cochlear model without computing the derivatives of $TF(\beta)$ as in (7).

Calculating the wave number $k(j\beta)$ and applying the definition (4b), we can compute $sZ_n(s)$ and then find a suitable rational approximation for it. Zweig's form for $sZ_n(s)$ derived from experimental data is $s^2+\delta\cdot s+1+\rho e^{-s\cdot\psi}$ [29]. This expression hints at a possible rational form for $sZ_n(s)$, which is more easily implementable in hardware. If we replace the pure delay term $\rho e^{-s\cdot\psi}$ with an all-pass rational Pade approximant, $sZ_n(s)$ becomes a ratio of polynomials, with the order of the numerator polynomial higher than the order of the denominator polynomial by 2. The work [29] emphasized the importance of two conjugate zero pairs near $s=\pm j$ in $sZ_n(s)$ for providing collective amplification in the cochlea. Therefore, the simplest and most general rational form for $sZ_n(s)$ that approximates cochlear behavior in its low-intensity linear regime is given by:

$$s\cdot Z_n(s)=\frac{\left(s^2+2d\cdot s+1\right)^2}{s^2+s\frac{\mu}{Q}+\mu^2} \qquad (8)$$

Here, $1/2d$ is the quality factor of the two overlapping zero pairs in the CP impedance, while μ and Q are the natural frequency and quality factor of its pole pair. We show that the values of these parameters can be determined by matching desired features of the cochlear response.

3.2.1. Nonlinear Transmission-Line Model Formulation

Following [25], [32] and [33] we may write the local CP impedance in the form:

$$Z_n(s;\gamma)=Z_p(s)+\gamma\cdot Z_a(s) \qquad (9)$$

where $Z_a(s)$ represents the maximum contribution of the outer-hair-cell (OHC) active force generators, $Z_p(s)$ represents the "passive" cochlea, and the real parameter $0\leq\gamma\leq 1$ is the efficiency of the OHC transduction that depends only on the amplitude of local CP motion. In the low-intensity limit, γ approaches 1. At high levels, γ approaches 0. Its precise value at any given level depends on the form of OHC force nonlinearity, which is assumed to be memoryless and instantaneous, and was calculated in [25]. The parameter γ defines a family of linear models. According to the EQ-NL theorem [33], each model in the family has the same input-output cross-correlation function as a nonlinear model, where the input-output cross-correlation function is measured with flat-spectrum wideband noise at some input intensity. Such a wideband noise input enables one value of γ to characterize OHC saturation throughout the cochlea as all parts of it are equally stimulated and equally saturated. The substitution of the analysis of one nonlinear model with the analysis of a family of linear models is valid only for wideband noise stimuli. The response of the nonlinear model to a single tone stimulus will be quite different. Nevertheless, a linear analysis with □ is useful for designing a cochlea with a local nonlinearity (whether due to a slowly varying automatic gain control (AGC) or instantaneous function) such that its impedances have a well-characterized behavior for one class of inputs that is predictable from theory. Cochlear responses to other classes of inputs then emerge from the designed impedances and may be verified to yield desirable properties through experiments or simulations.

Two impedances define the model: The "passive" impedance $Z_p(s)$, and the low-level "threshold" impedance $Z_n(s;1)$. A simple resonator is usually chosen to represent the passive CP such that

$$s \cdot Z_p(s) = s^2 + s \cdot \delta_0 + 1 \tag{10}$$

Here, δ_0 is the damping when the OHC is disabled. Equation (8) represents the low-level linear limit corresponding to □$=1$:

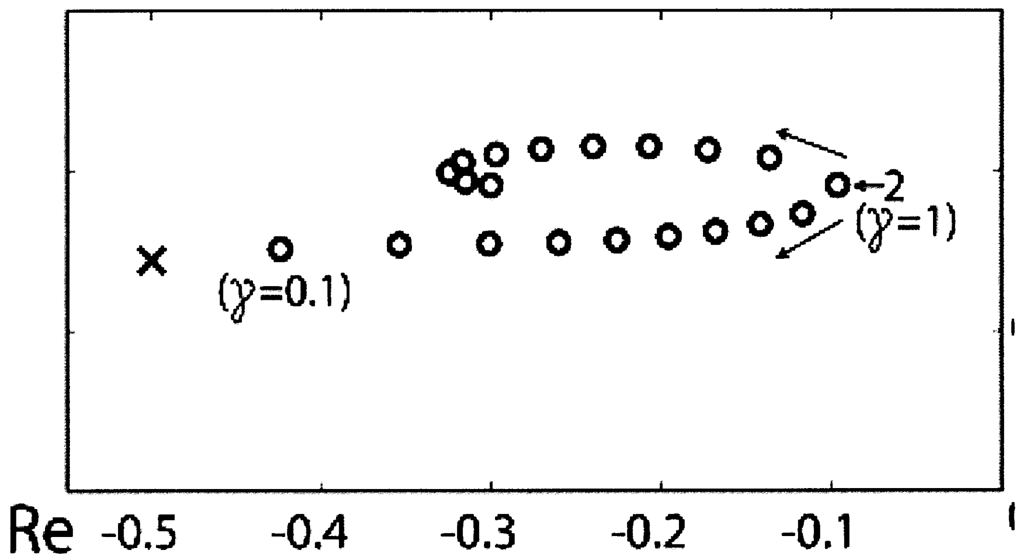
$$s \cdot Z_n(s;1) = \frac{\left(s^2 + s \cdot f_z / Q_z + f_z^2\right)^2}{s^2 + s \cdot \zeta_p + f_p^2} \tag{11}$$

where f_z and Q_z characterize the double-zero, and f_p and ζ_p characterize the pole location of the shunt impedance. Some simple but tedious algebra and (9), (10), and (11) yield an expression for the maximum contribution of the OHC active force generators $Z_a(s)$ given by

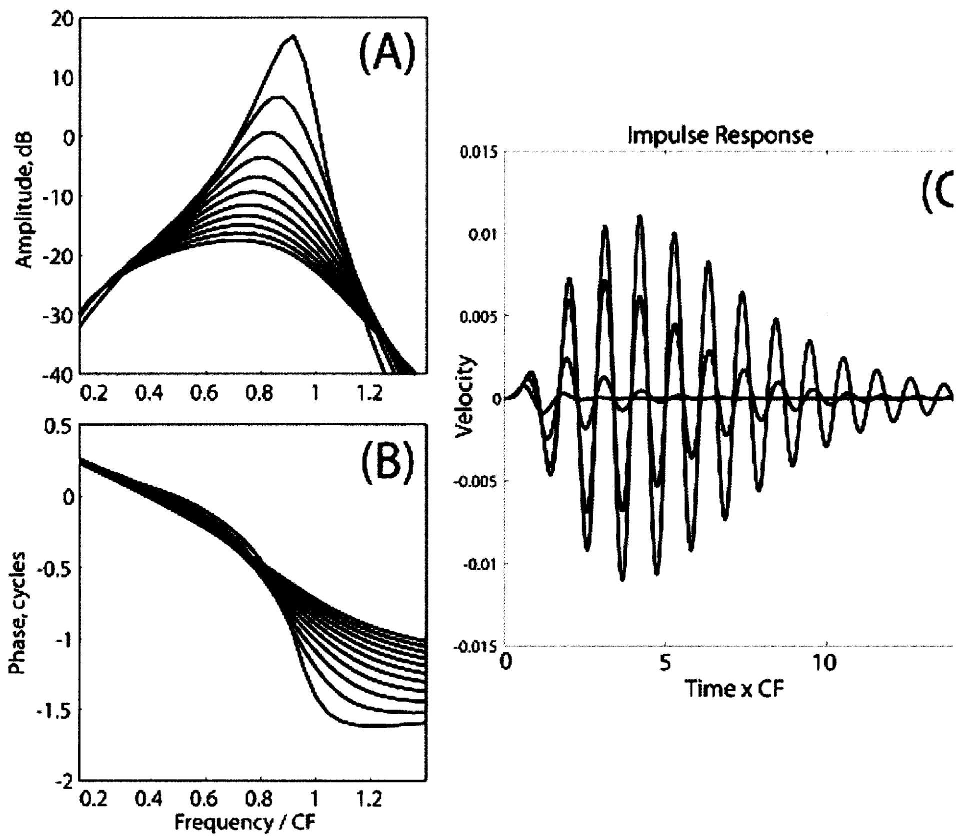
$$Z_a(s) = \delta - \delta_0 + \rho \frac{s^2 + s \cdot a_1 + a_0}{s \cdot \left(s^2 + s \cdot \zeta_p + f_p^2\right)} \tag{12}$$

where δ, ρ, a_1 and a_0 are computed from f_z, Q_z, f_p and ζ_p. The parameter δ represents the effective damping with maximum OHC contribution.

It was conjectured in [31] that requiring the near-invariance of fine time structure of the cochlear impulse response with stimulus intensity, an experimentally observed fact **[2]**, implies that the zeros of the effective local CP impedance $Z_n(s;\gamma)$ move almost perpendicularly to the $j\omega$-axis as the parameter γ is varied. This requirement imposes some constraints on the values of our parameters. Our system continuously moves from having two zero-pairs and one pole-pair to having only one zero-pair as the parameter γ is reduced from 1 to 0. Therefore, the double-zero must have separated somewhat, with one zero pair moving into the pole pair and getting cancelled, and the other zero pair becoming the "passive" CP impedance $Z_p(s)$. Requiring the two zero pairs to move almost perpendicular to the $j\omega$-axis implies that the two zero pairs and pole pair of $Z_n(s;1)$ and the zero pair of $Z_p(s)$ are nearly on a line that is perpendicular to the $j\omega$-axis. $Z_p(s)$ defines that line, and its damping parameter δ is picked to yield a desirable passive response. Fixing Q_z to yield a desirable active response then defines f_z, and choosing ζ_p to mimic a shift in the peak frequency as γ goes from 1 to 0 defines f_p. Values of δ, ρ, a_1 and a_0 are then computed.



[3]
[4] Fig. 1. Trajectories of the zeros of $sZ_n(s;\gamma)$ as a function of γ (only $Im>0$ part of each complex conjugate pair is shown). At $\gamma=1$ $sZ_n(s;\gamma)$ has two overlapping conjugate zero pairs and a conjugate pole pair. As the signal level increases, γ decreases, the overlapping zero pairs separate slightly and create two non-overlapping conjugate zero pairs; the zeros move out nearly perpendicularly to the frequency axis, as shown with the arrows; the conjugate pole pair does not move with γ. We show zero positions for $0\leq\gamma\leq1$, in steps of 0.1 in γ. As γ decreases to 0, one of the zero pairs cancels the pole pair, and $sZ_n(s;\gamma)$ becomes a passive $sZ_p(s)$ with only one zero pair.



[1]
[2] Fig. 2. Intensity dependence of cochlear transfer functions and impulse responses simulated with 16 sections per octave over 6 octaves (96 sections overall): (A) amplitude in dB; (B) phase in cycles (one cycle corresponds to 2π radians); (C) impulse response. Cochlear transfer functions $TF(\beta;\gamma)$ are measured as cross-correlation functions with wide band noise as an input stimulus, as γ varies from 0 to 1 in steps of 0.1. The abscissa is the dimensionless frequency variable $\beta=f/f_{CF}(x)$. Impulse responses are computed using inverse Fourier transforms of those transfer functions at several values of $\gamma=\{1, 0.9, 0.7, 0\}$. The abscissa is the dimensionless time variable $\tau=t\cdot f_{CF}(x)$.

3.2.2. Results

The discrete cochlea was simulated with 16 sections per octave over 6 octaves (96 sections overall). The parameter values used were $\delta_0=0.6$, $Q_z=5$, $\zeta_p=1$, $f_p\approx0.88$, $f_z\approx0.96$, yielding $\delta\approx-0.6$, $\rho\approx0.7$, $a_1\approx-0.24$, and $a_0\approx0.1$.

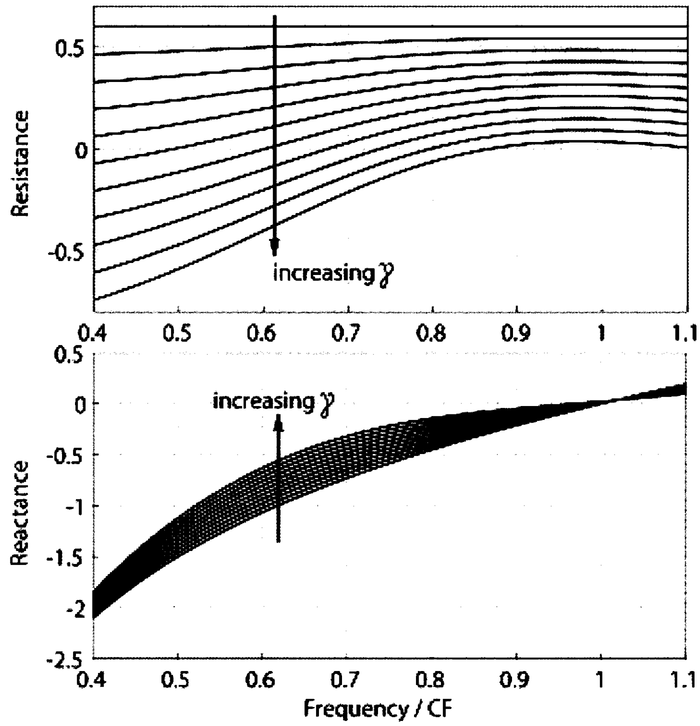
Parasitic reflections from section-to-section discontinuities may be observed by examining frequency responses: These reflections cause the standing wave in a basal region of the cochlea to manifest as a series of notches in the transfer function amplitude and as a series of jumps in the transfer function phase. Through simulations, we found that these reflections are avoided if $N\leq1.3$ such that the phase accumulation over a section is not too large.

Figure 1 shows the trajectories of the zeros of $sZ_n(s;\gamma)$ as γ varies from 1 to 0 in steps of 0.1. As expected, the double zero pair separates slightly with one zero pair moving towards a pole pair to eventually get cancelled and the other zero pair moving towards the passive-impedance zero pair location. All zeros move out nearly perpendicularly to the frequency axis.

Figures 2 (A) and (B) show cochlear transfer functions $TF(\beta;\gamma)$, computed using γ as a parameter, with $0\leq\gamma\leq1$, varying in steps of 0.1. The transfer functions $TF(\beta;\gamma)$ at different γ are input-output cross-correlation functions obtained from a nonlinear model with wideband white noise at different intensities as an input stimulus. Cochlear amplification, defined as the ratio of peak gains in the low-level linear limit to that of the passive cochlea, is 34 dB. The maximum Q_{-10dB} of the transfer function is 4.7, and the high-frequency slope is about 146 dB/octave.

Figure 2 (C) shows impulse responses, normalized by input at the stapes, and computed using inverse Fourier transforms of $TF(\beta;\gamma)$ at several values of $\gamma=\{1, 0.95, 0.7, 0\}$. As the intensity of the input stimulus increases and γ decreases, the envelope of the impulse response decreases in maximum amplitude and in duration, peaking at earlier times, but fine time structure remains nearly invariant, as can be seen from the timing of zero crossings of the response.

Figure 3 shows local CP impedances $Z_n(j\beta;\gamma)$. The top panel shows resistance $Re\{Z_n(j\beta;\gamma)\}$ and the bottom panel shows reactance $Im\{Z_n(j\beta;\gamma)\}$ as γ varies from 0 to 1 in steps of 0.1. Although the impedances are calculated for a given CP section for various frequencies, we can also interpret Figure 3 as showing local CP impedances along the length of the CP for a fixed frequency f and fixed



[5]
[6] Fig. 3. Intensity dependence of local CP impedances $Z_n(j\beta;\gamma)$. The top panel shows resistance $Re\{Z_n(j\beta;\gamma)\}$ and the bottom panel shows reactance $Im\{Z_n(j\beta;\gamma)\}$ as γ varies from 0 to 1 in steps of 0.1. The abscissa is the dimensionless frequency variable $\beta=f/f_{CF}(x)$. The discrete cochlea was simulated with 16 sections per octave over 6 octaves (96 sections overall).

parameter γ, and varying with the position x as determined from $\beta=f/f_{CF}(x)$. At the low-level limit the real component of the impedance is negative basal to the place where the transfer function amplitude peaks, and turns positive apical to that place. A negative real component of the local impedance indicates energy transfer to the traveling wave, while positive resistance indicates absorption of energy from it. The imaginary component of the impedance is negative basal to the place of local resonance, defined to be the place at which the imaginary component of $Z_n(s)$ is zero. The transfer function peaks at a place basal to the local resonance. The magnitude of the imaginary component of the impedance is higher than that of the real component; therefore, it plays a larger role in local CP impedance magnitude.

The effect of the double zero in $Z_n(s)$ is to decrease the magnitude of the local CP reactance making it close to zero over an extended region basal and around the place of transfer-function peaking such that more stages can provide significant cochlear amplification. A single zero is not as effective as a double zero because the double zero has a more extended region of frequencies (or places) where it is small [29].

The decrease in the magnitude of the local CP impedance allows high CP velocities to be achieved without a significant increase in pressure difference and associated high infusions of energy into the traveling wave. As the intensity of the input stimulus increases and γ decreases, the CP resistance becomes less negative, and the place where it crosses to positive values moves closer to the base. This collective behavior causes a shift in the peak frequency as the input intensity varies [31]. The decrease in CP reactance magnitude diminishes as well, making it less close to zero, especially basal to the new peak in transfer-function amplitude. The reduced negative resistance and increased local CP impedance magnitude reduce the height of the transfer-function amplitude peak as the input level increases. At $\gamma \approx 0.45$ the CP resistance becomes positive everywhere, but the cochlear amplifier still helps in reducing net local CP resistance, thereby increasing cochlear gain.

3.3. Time-Domain Stability Analysis

Implementing our cochlear model in hardware requires that the system be bounded-input bounded-output (BIBO) stable. A technique to determine the stability of time-domain solutions from frequency-domain transfer functions was proposed in [36] for linear active cochlear models. However, this technique is based on comparison of two numerically computed functions, and therefore can not unambiguously determine whether the system is stable or not. In our cochlear model the local CP impedance $Z_n(s)$ is rational, so the system has a finite order. Transforming (1a) and (1b) with rational $Z_n(s)$ in (8) or (11) into the time domain and discretizing in x using equally spaced spatial mesh to reflect the known nature of the cochlear response, we apply a state-space representation method to investigate BIBO stability:

$$\frac{d}{dt}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}u_{in} \tag{13}$$

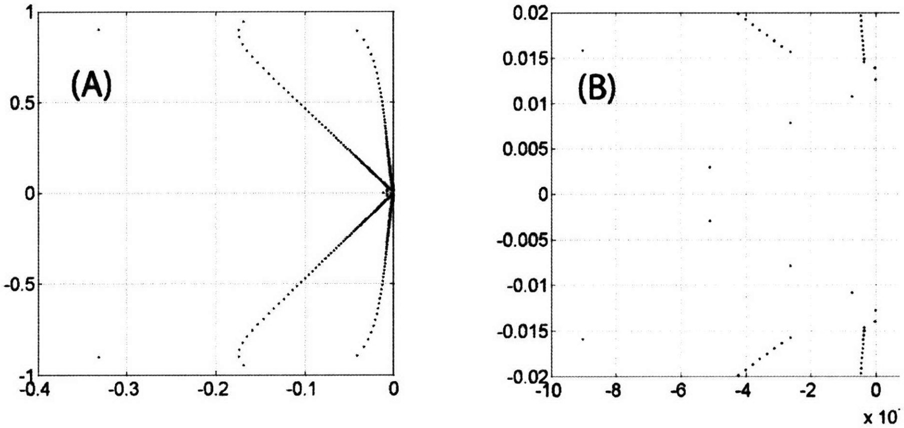
where $\mathbf{x}$ is a vector of state variables such as currents in inductors and voltages on capacitors, u_{in} is an input scalar signal, $\mathbf{B}$ is a column vector, and $\mathbf{A}$ is the state-space matrix of our cochlear model determined by the spatial discretization and the parameters describing $Z_n(s)$. The necessary and sufficient condition for BIBO stability of the system is that all the eigenvalues of $\mathbf{A}$ have negative real parts:

$$\mathrm{Re}\{\lambda_i(\mathbf{A})\} < 0 \tag{14}$$

Criterion (14) ensures that hardware or software implementations of the cochlear model will be stable.

Figure 4 (A) shows all the eigenvalues of matrix **A** for the cochlear model with $\gamma=1$ with 23 sections per octave over 6 octaves (138 sections overall). The termination of this system at the apical end is discussed later. Parasitic reflections from section-to-section discontinuities are insignificant if $N\leq 2.2$. The cochlear amplification is 45dB. Figure 4 (B) shows low-frequency eigenvalues in more detail. Note that the biggest challenge to the stability of the system comes from the eigenvalues with a frequency of approximately $0.9\cdot\omega_c(0)\cdot 2^{-6}\approx 0.014\cdot\omega_c(0)$. An input tone of that frequency causes the response to peak at the apical termination. This peaking suggests how instability arises in the system.

Suppose there is a weak low-frequency signal somewhere in the cochlea propagating toward the apex. It propagates without appreciable attenuation and then undergoes significant cochlear amplification just basal to the place of its peak. A large reflected wave is produced if the amplitude of this signal peaks at the apical termination, and the termination impedance deviates from the characteristic impedance at the apex of the cochlea around the frequency of this signal. This reflected signal undergoes still more amplification just basal to the place of its peak, because it travels back through the region where the local CP resistance is negative and the reactance magnitude is small. This signal now propagates back toward the base of the cochlea without appreciable attenuation, and reflects from basal or section-to-section discontinuities, to create a return signal and enable multiple reflections. If cochlear amplification is significant, and the terminating impedance at the apex does not match the characteristic impedance precisely, the net round-trip gain can become larger than 1, causing instability due to buildup of reflections.



[7]
[8] Fig. 4. (A) Eigenvalue plot for $\gamma=1$ with 23 sections per octave over 6 octaves (138 sections overall). The model is stable if the real parts of all the eigenvalues are negative. The cochlear amplification is 45dB. (B) A close-up view of the low-frequency eigenvalues, which present the biggest challenge to the stability of the system due to reflections from mismatched apical termination. The termination consists of a resistor and an inductor in series approximating the characteristic impedance at the apex at frequencies both higher and lower than the CF at the apex. The system is barely stable. If the amplification is increased, instability results. Better apical termination is required in order to further increase cochlear amplification without going unstable.

Stabilizing a cochlea demands either improving the matching between the impedance of the apical termination section and the characteristic impedance of the cochlea at the apex, or reducing the amplification of the cochlear sections near the apex gradually, or both. Another technique of finding optimal initial conditions in the cochlea to minimize the effects of instability was presented in [37]. However, this technique is not suitable for hardware implementations since the initial conditions that were proposed are not easily controlled.

We first consider the technique of matching the impedance of the apical termination section to the characteristic impedance of the cochlea at the apex. Cochlear characteristic impedance was studied in [38]. The characteristic impedance Z_c in the cochlea is defined in terms of the volume velocity U and the pressure difference P:

$$\frac{1}{Z_c} \equiv \frac{U}{P} \tag{15}$$

The volume velocity U is proportional to dP/ds, as can be seen from the macromechanical equation (1a) rewritten in terms of variable s defined by (3). We then obtain:

$$\frac{1}{Z_c(s)} = -\frac{1}{l \cdot L(0) \cdot \omega_c(0)} \cdot \frac{d}{ds} \ln\left[P(s)\right] \tag{16}$$

Substituting the WKB-approximate solution for the pressure difference $P(s)$ from (5) into (16) yields the following expression for the characteristic impedance of the cochlea:

$$\frac{1}{Z_c(s)} = \frac{k(s)}{l\cdot L(0)\cdot\omega_c(0)}\cdot\left[1-\frac{1}{2}\cdot\frac{d}{ds}\left(\frac{1}{k(s)}\right)\right] \tag{17}$$

Once again, assuming that the properties of the cochlea scale slowly relative to the wavelength of the traveling wave, i.e., $\left|\frac{d}{ds}\left(\frac{1}{k(s)}\right)\right| \ll 1$, we can ignore the second term in the brackets that comes from the pre-exponential term in (5). We obtain:

$$\frac{1}{Z_c(s)} \approx \frac{k(s)}{l\cdot L(0)\cdot\omega_c(0)} \tag{18}$$

Invoking the definition (4b), we can rewrite (17) and (18):

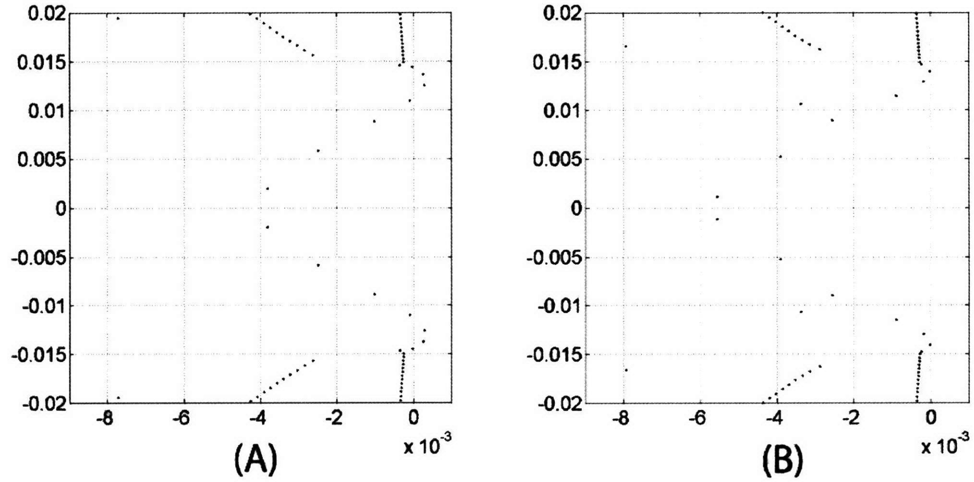
$$\frac{\omega_c(0)\cdot l\cdot L(0)}{Z_c(s)} = \frac{4N}{\sqrt{sZ_n(s)}} - \frac{1}{4}\cdot\frac{\frac{d}{ds}\left[sZ_n(s)\right]}{sZ_n(s)} \approx \frac{4N}{\sqrt{sZ_n(s)}} \tag{19}$$

It is evident that the form of the cochlear characteristic impedance $Z_c(s)$ might be irrational even if the CP impedance $Z_n(s)$ is rational. Therefore, we use a rational approximation to synthesize the admittance $G_t(j\omega)$ that terminates the cochlea at the apical end:

$$G_t(j\omega) \approx \frac{1}{Z_c\left(\frac{j\omega}{\omega_c(x_0)}\right)} \tag{20}$$

We design $G_t(j\omega)$ to achieve good impedance match at the frequencies around $\omega_c(x_0)$, where x_0 is the location of the apical termination, and $\omega_c(x_0) = \omega_c(0)\cdot exp(-x_0/l)$ is the CF at the apex. It is also immediately obvious from (19) that the characteristic impedance of the cochlea depends on the signal amplitude through the parameter γ. This dependence should be taken into account when designing the apical termination. In the simplest case, the admittance $G_t(j\omega)$ is designed for the low-level limit $\gamma=1$ in the hope that even as γ decreases with increasing signal level, and the deviation of the admittance $G_t(j\omega)$ from the cochlear characteristic admittance at the apex grows along with the reflection coefficient, the reduction in cochlear amplification associated with lower γ will still reduce the reflected signal and therefore decrease the round-trip gain. In this case, stability at $\gamma=1$ also guarantees stability for $\gamma<1$.

One of the simplest rational forms for the termination is a resistor and an inductor in series approximating the characteristic impedance at the apex at frequencies both higher and lower than



[9]
[10] Fig. 5. A comparison of cochlear stability for two types of apical termination impedance: (A) A resistor and inductor in series approximate the characteristic impedance at the apex at frequencies both higher and lower than the CF at the apex but the system is unstable at 51dB of amplification. (B) An impedance described by the ratio of a fourth-order polynomial to a third-order polynomial approximates the characteristic impedance at the CF at the apex more accurately and leads to a barely stable system at 51dB of amplification. The cochlear model has 30 sections per octave over 6 octaves (180 sections overall). If the amplification is increased beyond 51dB, instability results, and a better termination technique becomes necessary.

$\omega_c(x_0)$. At frequencies lower than $\omega_c(x_0)$ the resistor dominates, and its resistance can be computed from (20), (19) and (8):

$$R_t \approx \frac{\omega_c(0)\cdot l\cdot L(0)}{4N\cdot\mu} \tag{21}$$

At frequencies higher than $\omega_c(x_0)$ the inductor dominates, and its inductance can be computed from (20), (19) and (8):

$$L_t \approx \frac{l\cdot L(x_0)}{4N} \tag{22}$$

where $L(x_0)= L(0)\cdot exp(x_0/l)$ is the per-length fluid mass in the scalae at the apex.

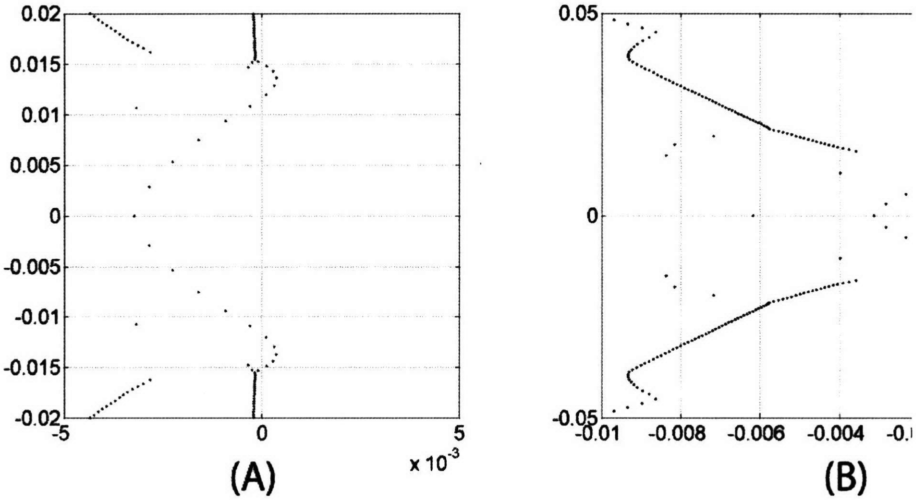
The simple apical termination defined by (21) and (22) does not depend on γ, but nonetheless ensures the stability of the system for cochlear amplifications no higher than 45dB. This level of amplification is achieved with 23 sections per octave and $N=2.2$, as shown in Figure 4 (A) and (B). However, the system is barely stable. If the amplification is increased above 45dB, instability results. A better apical termination is required in order to further increase the cochlear amplification without going unstable.

Figure 5 (A) shows low-frequency eigenvalues of matrix **A** for the cochlear model with 30 sections per octave over 6 octaves (180 sections overall). The simple apical termination defined by (21) and (22) is used. The standing wave in the basal region of the cochlea due to the parasitic reflections from section-to-section discontinuities is avoided if $N \leq 3.0$. The cochlear amplification is 51dB. However, the system is clearly unstable, and increasing the number of sections per octave beyond 30 leads to just more populated plot that looks like Figure 5 (A), but does not improve the stability. A more sophisticated apical termination is required to achieve BIBO stability in the cochlear model. Figure 5 (B) shows the same cochlea when a higher-order approximation to the characteristic impedance at the apex is used to terminate it:

$$G_t\left(s = \frac{j\omega}{\omega_c(x_0)}\right) = \frac{4N}{\omega_c(0)\cdot l\cdot L(0)}\cdot\frac{c_0 + c_1 s + c_2 s^2 + c_3 s^3}{1 + d_1 s + d_2 s^2 + d_3 s^3 + d_4 s^4} \tag{23}$$

The eight parameters in (23) were chosen to accurately approximate the cochlear characteristic impedance, calculated in (19), for $\gamma=1$. Two degrees of freedom were used to approximate the characteristic impedance at frequencies much lower and much higher than the CF at the apex. The six other degrees of freedom were used to accurately approximate the characteristic impedance at approximately $0.9\cdot\omega_c(x_0)$, which is where peaking occurs at the apex. Note that the apical termination defined in (23) does not depend on γ. This termination ensures the stability of the system with cochlear amplification of up to 51dB, which extends the use of the single-section termination technique by 6dB. If the amplification is increased above 51dB, the cochlea with this single-section termination becomes unstable, and even better apical termination techniques become necessary.

The perfectly matched layer (PML) concept originally devised in [39] involves surrounding the computational domain with an artificial layer which absorbs outward traveling electromagnetic waves, thus preventing parasitic reflections into the computational domain. Several related techniques gradually reduce the amplification of the cochlear sections towards the apex, and were described in [40], [41]. In one of them the damping of the simple resonator that characterizes the CP impedance in the passive cochlear model was gradually increased towards the apex [40]. This technique is not applicable for the models presented in this section since these models are active, and the local CP impedance is not a simple resonator. Another technique gradually introduces viscosity into longitudinal fluid coupling between the cochlear sections near the apex [41]. This scheme reduces the amplification of the cochlear sections helping to stabilize the system. Although this method could be used in the models presented in this section, it would likely require redesigning the sections near the



[11]
[12] Fig. 6. A comparison between two techniques for apical termination in a cochlear model with 50 sections per octave over 6 octaves (300 sections overall): (A) Termination with an impedance implemented as the ratio of a fourth-order polynomial to a third- order polynomial approximates the characteristic impedance at the apex but the system is highly unstable at 72dB of amplification. (B) Termination implemented by gradually decreasing the gains of stages towards the apex; gain tapering is accomplished by lowering Q_z from 5 to 5/1.5 over the last 75 sections (1.5 octaves). The tapering technique reduces the strength of reflected signals instead of reducing the apical reflection coefficient. The tapering needs to be gentle enough to avoid reflections from the associated discontinuities. The system is stable when the first 4.5 octaves have 72dB of amplification.

apex in an actual hardware implementation. In contrast, we reduce the amplification of the cochlear sections near the apex by gradually decreasing Q_z towards the apex. Such a technique is easily realizable in hardware, does not require redesigning the sections near the apex, and appears to be a solution seen in biology as well.

Figure 6 (A) shows low-frequency eigenvalues of **A** for the cochlear model with 50 sections per octave over 6 octaves (300 sections overall). The single-section apical termination defined by (23) was used. Parasitic reflections from section-to-section discontinuities are insignificant if $N \leq 5.24$. The cochlear amplification is 72dB. The system is unstable, and increasing the number of sections per octave does not improve the stability. The single-section apical termination is not robust enough to ensure BIBO stability. Figure 6 (B) shows the same cochlear model using the apical termination technique of gradually decreasing the gains of stages towards the apex. Gain tapering is accomplished by lowering Q_z from 5 to 5/1.5 over the last 75 sections (1.5 octaves). The tapering needs be gentle enough to avoid reflections from the associated discontinuities. Figure 6 (B) reveals a second set of eigenvalues that challenge the stability of the system. These occur at a frequency of approximately $0.9\cdot\omega_c(0)\cdot 2^{-6}\cdot 2^{1.5}\approx 0.04\cdot\omega_c(0)$. An input tone of that frequency causes the cochlear response to peak at the place where we start tapering the gain. The system is nevertheless stable, and achieves a cochlear amplification of as high as 72dB in the first 4.5 octaves, indicating the superiority of this technique over single-section matching techniques. The price for the increased robustness of the gain tapering technique is an increase of chip area or reduced cochlear amplification at the low frequencies when the system is implemented in hardware.

3.4. Composite Cochlear Architecture

3.4.1. Analysis

The biological cochlea appears to primarily support forward traveling waves [34] such that its architecture can be simplified into a unidirectional filter cascade, an architecture that we shall term composite in keeping with prior convention [42]. Determining the transfer functions of the filters in the cascade yields simple models suitable for hardware implementation. A second-order low-pass filter section was proposed in [43], and an analog low-power wide-dynamic-range integrated-circuit cochlea was built [6] with such composite architectures. However, the low-pass resonant filter sections in these cochleae gave rise to excessive group delay of the cochlear output when compared with auditory data [44]. A compromise between the desired frequency response and the group delay was achieved in [44] by introducing a zero pair and increasing the number of poles in the filter section to three. In addition, practical realization of such a composite cochlea required additional second-order output filters to enhance resolution [42]. The modified transmission-line model in [42] also uses a filter cascade, but with the filter transfer functions now representing isolated sections of a one-dimensional transmission line with simple mass-elasticity-damping local impedances. Each section of such a filter cascade is formed by isolating a lumped section of a one-dimensional transmission line from its neighbors and loading it with the characteristic impedance of the cochlea in that location. The transfer function of a filter section that represents a one-dimensional transmission line with simple mass-elasticity-damping shunt impedances was derived in [45]. This stage has a relatively high-Q zero pair and a pole pair combined with a low-Q zero pair and pole pair. A second-order approximation to this transfer function, formed by dropping the low-Q zero and pole pairs and adjusting the high-Q zero and pole pairs, was implemented in [42]. However, relatively high values of Q and additional second-order output filters were still needed [42]. In this section we derive transfer functions for filters in a composite cochlea from our transmission-line model with the local impedance (8).

We start with the WKB-approximate analytical solution for the cochlear TF (6). Recognizing that the essence of cochlear action is collective amplification, represented by the exponential term in (6), and, therefore ignoring the pre-exponential dependencies, *TF(s)* can be written as:

$$TF(s) \propto \prod_{i=1}^{m} \exp\left(-\int_{s_{i-1}}^{s_i} k(s')\,ds' \right) \tag{24}$$

We have split up the integral in the exponential into m smaller regions of integration (with $s_0=0$ and $s_m=s$). If the input to the system is a pure tone with fixed frequency ω, then $s_i=j\omega/\omega_c(x_i)$, and (24) describes spatial propagation of the signal in a cascade of filter stages, with the transfer function of the i-th stage being

$$H_i = \exp\left(-\int_{s_{i-1}}^{s_i} k(s')\,ds' \right) \tag{25}$$

where x_{i-1} is the location of the input of the i-th stage (and the output of the $i-1$ th stage), and x_i is the location of the output of the i-th stage (and the input of the $i+1$ th stage). If this composite cochlea is finely quantized, i.e., enough stages are used, s_i-s_{i-1} and $\int_{s_{i-1}}^{s_i} k(s')\,ds'$ become small. Using the approximation $exp(-x)\approx 1/(1+x)$ for $x \ll 1$, H_i can be written as:

$$H_i \approx \frac{1}{1+\int_{s_{i-1}}^{s_i} k(s')\,ds'} \approx \frac{1}{1+k(s_i)\cdot(s_i - s_{i-1})} \tag{26}$$

where we have assumed that $k(s')\approx k(s_i)$ over the (small) interval $[s_{i-1}, s_i]$.

As in sections II and III, the discretization of the model in x is chosen to reflect the known nature of the cochlear response. The most efficient implementation of the algorithm uses a spatial mesh that is equally spaced in x, resulting in an exponential taper of filter characteristic frequencies in the cascade. Therefore, we have:

$$\frac{s_{i-1}}{s_i} \equiv \frac{\omega_c(x_i)}{\omega_c(x_{i-1})} = e^{-\frac{\Delta x}{l}} = 2^{-\frac{1}{N_{oct}}} \tag{27}$$

where $\Delta x \equiv x_i - x_{i-1}$ is the constant length of the interval in spatial quantization, and N_{oct} is the number of filters per octave span. Using (27) and (4b), we can rewrite (26) as:

$$H_i = H(s_i) = \frac{1}{1 + \dfrac{\alpha \cdot s_i}{\sqrt{s_i Z_n(s_i)}}} \tag{28}$$

where $H(s)$ is the normalized filter transfer function and α is a constant given by

$$\alpha \equiv 4N \cdot \left(1 - 2^{-\frac{1}{N_{oct}}}\right) \approx \frac{4N \cdot \ln(2)}{N_{oct}} \tag{29}$$

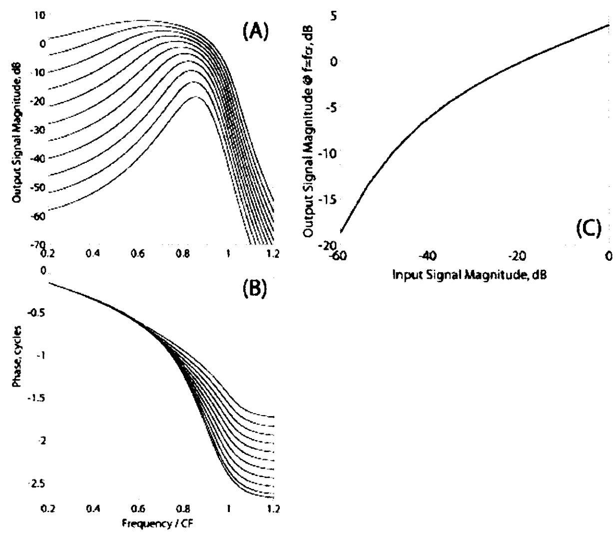
Note that (28) and (29) could alternatively be obtained by isolating each lumped section of the one-dimensional transmission line from its neighbors and loading it by the cochlear characteristic impedance $Z_c(s_i)$, as in [42]. Ignoring the local admittance with respect to the characteristic admittance then yields a voltage divider between the longitudinal fluid coupling impedance $j\omega \cdot L(x_i) \cdot \Delta x = s_i \cdot \omega_c(0) \cdot L(0) \cdot \Delta x$ and the characteristic impedance $Z_c(s_i)$ computed in (19). The transfer function of this divider again gives (28) and (29).

Using $sZ_n(s)$ in the form of (8), we obtain:

$$H(s_i) = \frac{1}{1 + \dfrac{\alpha \cdot s_i \cdot \sqrt{s_i^2 + s_i \cdot \mu / Q + \mu^2}}{s_i^2 + 2 \cdot d \cdot s_i + 1}} \tag{30}$$

The expression in (30) is not a rational function in s_i, and therefore cannot be implemented using a lumped system. However, the magnitude and phase-response shapes of the model defined by (8) are not sensitive to the value of Q. Setting $Q=0.5$ completes a square under the square root, and the normalized filter transfer function becomes rational:

$$H(s_i) = \frac{1}{1 + \dfrac{\alpha \cdot s_i \cdot (s_i + \mu)}{s_i^2 + 2 \cdot d \cdot s_i + 1}} = \frac{s_i^2 + 2 \cdot d \cdot s_i + 1}{(1+\alpha) s_i^2 + (2 \cdot d + \alpha \cdot \mu) \cdot s_i + 1} \tag{31}$$



[13]

[14] Fig. 7. Frequency response of the nonlinear composite cochlea measured at position x_0 to a *pure tone* input stimulus with frequency f and amplitudes varying from -60dB to 0dB in steps of 6dB. The abscissa is the dimensionless frequency variable $\beta=f/f_{CF}(x)$: (A) amplitude of the output signal in dB; (B) phase in cycles (one cycle corresponds to 2π radians). (C) Output amplitude as a function of the input-stimulus amplitude at the small-signal peak frequency. The composite cochlea was simulated with 25 sections per octave over 6 octaves (150 sections overall).

The pole pair at $\beta_p \approx 1/\sqrt{1+\alpha} < 1$ causes the amplitude of the normalized transfer function (31) of the section to first peak, and then drop sharply due to the zero pair at slightly higher frequency $\beta_z \approx 1$. The peak contributes to the collective cochlear amplification, while the near-null sharpens the roll-off slope beyond the peak and thus increases frequency resolution. The zero pair also nearly offsets the group delay accumulation due to the pole pair of the filter.

To preserve the nonlinear properties of the transmission-line model in section II, we require that only the damping d of the local CP impedance zeros in (8) varies with the signal level. If d^2□ 1, the double-zero pair in (8) moves almost perpendicularly to the $j\omega$-axis similarly to Fig. 1. Therefore, d in (31) depends only on the envelope of the local signal, i.e., either the input or output of the i-th filter in the cascade, which simplifies the design of the AGC. In this work we simulate a linear dependence of d on filter's input signal envelope $|A|$, the "power-1 law" nonlinearity also implemented in [6]: $d(|A|)=d_{min}+\sigma\cdot|A|$.

As the parameter μ in (31) increases, the shift in the peak frequency with the stimulus level increases, but the collective cochlear amplification drops as the poles of each filter section become more and more over-damped. As the parameter α in (31) increases, cochlear amplification grows, but so does the group delay. This represents a tradeoff between group delay and amplification in the cochlea.

3.4.2. Results

A composite cochlea was simulated with $N_{oct}=25$ sections per octave over 6 octaves of frequency (150 sections overall). The parameter values $d_{min}=0.1$, $d_{max}=0.5$, $\mu=0.45$, $\alpha\approx0.33$ were chosen so that the cochlear responses $TF(\beta;d)$ measured as cross-correlation functions with a wide band noise as an input stimulus, when parameter d is varied linearly from d_{min} to d_{max}, are close to those of the transmission-line cochlear model in section II shown in Fig. 2 (A), (B). The simulated cochlear amplification is 35 dB, the maximum Q_{-10dB} is 3.8, and the high-frequency slope is 200 dB/octave. Impulse responses, normalized by input at the stapes and computed as inverse Fourier transforms of $TF(\beta;d)$, are also close to those shown in Fig. 2 (C). The group delay in the composite architecture is slightly higher than in our transmission-line model, but still almost a factor of 3 lower than in [6]. The fine time structure invariance with stimulus intensity is satisfactory.

One of the most interesting characteristics of nonlinear cochlear implementations is the response to pure tone input stimuli of various frequencies and amplitudes. However, unlike with a wide band noise input stimulus, the shape of magnitude and phase responses to the pure tone stimulus depends on the form of the nonlinearity.

Figures 7 (A) and (B) show frequency responses of the nonlinear composite cochlea measured at position x_0 to pure tone input stimuli with frequencies f and amplitudes from -60dB to 0dB in steps of 6dB. The slope σ of the AGC function is chosen so that $d(|A|=0dB)=d_{max}$. Figure 7 (A) shows the amplitude of the output signal, Figure 7 (B) shows phase; Figure 7 (C) shows the output amplitude versus input amplitude for a pure-tone stimulus at the small-signal peak frequency. The curves of Figure 7 are very similar to those measured in the biological cochlea [1], [46].

3.5. Conclusion

We showed how single-mode transmission-line cochlear models can be efficiently implemented using rational approximations to the CP impedance with nonlinear active elements. We introduced the state-space representation method to analyze BIBO stability of the active cochlear model. The overall stability is shown to be greatly influenced by the amount of gain and the terminating impedance at the end of the cochlea. We derived an efficient composite model from transmission-line cochlea using a

WKB approximation. Each second-order filter in the cascade uses a simple AGC and the model exhibits many linear and nonlinear properties of the biological cochlea, including low group delay accumulation, steep roll-off and high resolution, and the fine time structure invariance with input signal level.

3.6. References

[1] M. Ruggero, N. Rich, A. Recio, S. Narayan, and L. Robles, "Basilar-membrane responses to tones at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.*, vol. 101, pp. 2151-2163, 1997.

[2] A. Recio, and W. Rhode, "Basilar membrane responses to broadband stimuli," *J. Acoust. Soc. Am.*, vol. 108, pp. 2281-2298, 2000.

[3] D. Johnson, "The relationship between spike rate and synchrony in responses of auditory-nerve fibers to single tones," *J. Acoust. Soc. Am.*, vol. 68, pp. 1115-1122, 1980.

[4] D. Louage, M. van der Heijden, and P. Joris, "Temporal properties of responses to broadband noise in the auditory nerve," *J. Neurophysiol.*, vol. 91, pp. 2051-2065, 2004.

[5] W. Liu, A. Andreou, and M. Goldstein, "Voiced-speech representation by an analog silicon model of the auditory periphery," *IEEE Trans. on Neural Networks*, vol. 3 (3), pp. 477-487, 1992.

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

[7] S. Zhak, S. Mandal, and R. Sarpeshkar, "A proposal for an RF cochlea," *in Proceedings of the Asia Pacific Microwave Conference*, Dec. 2004.

[8] R. Sarpeshkar, "Traveling waves versus bandpass filters: The silicon and biological cochlea," *Proc. of the Intl. Symp. Recent Developments in Auditory Mechanics*, H. Wada *et al.*, pp. 216—222, 2000.

[9] L. Turicchia, and R. Sarpeshkar, "A bio-inspired companding strategy for spectral enhancement," *IEEE Trans. Speech and Audio Proc.*, vol. 13 (2), 2005.

[10] R. Kumaresan and Y. Wang, "On representing signals using only timing information," *J. Acoust. Soc. Am.*, vol. 110, pp. 2421-2439, 2001.

[11] Z. Smith, B. Delgutte, A. Oxenham, "Chimaeric sounds reveal dichotomies in auditory perception," *Nature*, vol. 416, pp. 87-90, 2002.

[12] J. Rubinstein, B. Wilson, C. Finley, and P. Abbas, "Pseudospontaneous activity: stochastic independence of auditory nerve fibers with electrical stimulation," *Hear. Res.*, vol. 127, pp. 108-118, 1999.

[13] K. Nie, G. Stickney, and F. Zeng, "Encoding frequency modulation to improve cochlear implant performance in noise," *IEEE Transactions on Biomedical Engineering*, vol. 52(1), pp. 64-73, 2005.

[14] G. Zweig, "Auditory speech preprocessors," in *Proceedings of the DARPA Speech and Natural Language Workshop*, pp. 230-235, 1989.

[15] P. Loizou, "Mimicking the human ear," *IEEE Sig. Process. Mag.*, pp. 101-130, Sept. 1998.

[16] F. Spelman, "The past, present, and future of cochlear prostheses," *IEEE Engineering in Medicine and Biology*, p. 27, 1999.

[17] J. Allen, "Cochlear micromechanics – a physical model of transduction," *J. Acoust. Soc. Am.*, vol. 68, pp. 1660-1670, 1980.

[18] S. Neely, "A model of cochlear mechanics with outer hair cell motility," *J. Acoust. Soc. Am.*, vol. 94, pp. 137-146, 1993.

[19] A. Hubbard, "A traveling-wave amplifier model of the cochlea," *Science*, 259, pp. 68-71, 1993.

[20] F. Mammano, R. Nobili, "Biophysics of the cochlea: Linear approximation," *J. Acoust. Soc. Am.*, vol. 93, pp. 3320-3332, 1993.

[21] C. Geisler and C. Sang, "A cochlear model using feed-forward outer hair cell forces," *Hear. Res.*, vol. 86, pp. 132-146, 1995.

[22] R. Chadwick, E. Dimitriadis, and E. Iwasa, "Active control of waves in a cochlear model with subpartitions," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 93, pp. 2564-2569, 1996.

[23] T. Fukazawa, "A model of cochlear micromechanics," *Hear. Res.*, vol. 113, pp. 182-190, 1997.

[24] C. Steele, "Toward three-dimensional analysis of cochlear structure," *Otorhinolaryngol. Relat. Spec.*, vol. 61, pp. 238-251, 1999.

[25] E. de Boer and A. Nuttall, "The mechanical waveform of the basilar membrane. III. Intensity effects," *J. Acoust. Soc. Am.*, vol. 107, pp. 1497-1507, 2000.

[26] A. Hubbard, Z. Yang, L. Shatz, and D. Mountain, "Multi-mode cochlear models," *In* Symposium on Recent Developments in Auditory Mechanics. H. Wada *et al.*, Singapore, pp. 167-173, 2000.

[27] K. Lim and C. Steele, "A three-dimensional nonlinear active cochlear model analyzed by the WKB-numeric method," *Hear. Res.*, vol. 170, pp. 190-205, 2002.

[28] T. K. Lu, S. Zhak, P. Dallos, and R. Sarpeshkar, "Fast cochlear amplification with slow outer hair cells," *Hear. Res.*, vol. 214 (1-2), pp. 45-67, Apr. 2006.

[29] G. Zweig, "Finding the impedance of the organ of Corti," *J. Acoust. Soc. Am.*, vol. 89, pp. 1229-1254, 1991.

[30] P. Kolston, E. de Boer, M. Viergever, G. Smoorenburg, "What type of force does the cochlear amplifier produce?" *J. Acoust. Soc. Am.*, vol. 88, pp. 1794-1801, 1990.

[31] C. Shera, "Intensity-invariance of fine time structure in basilar-membrane click responses: Implications for cochlear mechanics," *J. Acoust. Soc. Am.*, vol. 110, pp. 332-348, 2001.

[32] S. Neely and D. Kim, "An active cochlear model showing sharp tuning and high sensitivity," *Hear. Res.*, 9(2), pp. 123-130, 1983.

[33] E. de Boer, "Connecting frequency selectivity and nonlinearity for models of the cochlea," *Aud. Neurosci.*, vol. 3, pp. 377-388, 1997.

[34] E. de Boer and E. van Bienema, "Solving cochlear mechanics problems with higher-order differential equations," *J. Acoust. Soc. Am.*, vol. 72, pp. 1427-1434, 1982.

[35] C. Shera, A. Tubis, and C. Talmadge, "Coherent reflection in a two-dimensional cochlea: Short-wave versus long-wave scattering in the generation of reflection-source otoacoustic emissions," *J. Acoust. Soc. Am.*, vol. 118, pp. 287-313, 2005.

[36] S. Koshigoe, A. Tubis, "Frequency-domain investigations of cochlear stability in the presence of active elements," *J. Acoust. Soc. Am.*, vol. 73, pp. 1244-1248, 1983.

[37] J. Xin, "Dispersive instability and its minimization in time-domain computation of steady-state responses of cochlear models," *J. Acoust. Soc. Am.*, vol. 115, pp. 2173-2177, 2004.

[38] S. Puria and J. Allen, "A parametric study of cochlear input impedance," *J. Acoust. Soc. Am.*, vol. 89, pp. 287-309, 1991.

[39] J.-P. Berenger, "A perfectly matched layer for the absorption of electromagnetic waves," *J. Comput. Phys.*, 114(1), pp. 185-200, 1994.

[40] M. van den Raadt and H. Duifhuis, "Different boundary conditions in a one-dimensional time domain cochlea model," *In* Biophysics of Hair Cell Sensory Systems. World Scientific, Singapore, p. 103, 1993.

[41] J. Xin, Y-Y. Qi, and L. Deng, "Time domain computation of a nonlinear nonlocal cochlear model with applications to multitone interaction in hearing," *Comm. Math. Sci.*, vol. 1(2), pp. 211-227, 2003.

[42] J. Kates, "Accurate tuning curves in a cochlear model," *IEEE Trans. Speech and Audio Proc.*, vol. 1(4), pp. 453-462, 1993.

[43] R. Lyon and C. Mead, "An analog electronic cochlea," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36(7), pp. 1119-1134, 1988.

[44] J. Eggermont, "Analysis of compound action potential responses to tone bursts in the human and guinea pig cochlea," *J. Acoust. Soc. Am.*, vol. 60(5), pp. 1132-1139, 1976.

[45] U. Eysholdt and V. Mellert, "A simulation of the motion of the basilar membrane using digital filters," *Acoustica*, vol. 33, pp. 116-121, 1975.

[46] N. Cooper and W. Rhode, "Mechanical responses to two-tone distortion products in the apical and basal turns of the mammalian cochlea," *J. Neurophysiol.*, vol. 78, pp. 261-270, 1997.

4. Multi-mode one-dimensional transmission-line cochlear models

***Abstract*-An approximate analytical technique for analyzing multi-mode one-dimensional transmission-line cochlear models is presented. This technique allows separating the modes, which can then be analyzed by any method for single-mode models, including the analytical WKB-approximate solution. The usage of this technique is illustrated by two examples: applying it to two-mode Sandwich, and to traveling-wave-amplifier cochlear models. The approximate analytical solutions agree qualitatively and quantitatively with the exact numerical simulations for both models. The presented technique also helps to provide physical intuition and insight into cochlear model functioning: In the two-mode Sandwich cochlear model, the effect of the additional transmission line is shown to be significant only in the cut-off region. In the traveling-wave-amplifier cochlear model the second transmission line is shown to be crucial for mimicking the behavior of the biological cochlea, such as high frequency resolution, large active amplification and steep roll-off. The second line lowers the impedance that the first line sees over an extended region basal to and around the peak, allowing high current peaks to be achieved without excessive infusion of energy into the traveling wave.**

4.1. INTRODUCTION

Mammalian cochlea provides frequency-to-location mapping with remarkable sensitivity, frequency resolution, amplification at the characteristic frequency (CF) and steep roll-off beyond CF in a broad frequency span of about 3 decades (10 octaves). It has an input dynamic range that spans 6 orders of magnitude in sound pressure. Numerous cochlear models of varying complexity have been proposed to account for these and many other features of the biological cochlea. The type of models, where two symmetric chambers of fluid are separated by a flexible membrane that consists of a number of sections coupled only by the fluid, received a lot of attention. Initially, the membrane was modeled as the simplest mass-elasticity-viscosity resonator with properties slowly varying along its length (Allen 1980, Watts 1993). Later works modeled the membrane as having an active process in the form of pressure sources controlled by the membrane motion. The longitudinal fluid coupling in the chambers can be approximated as one-dimensional (Kolston 1990, Zweig 1991, Neely 1993, Fukazawa 1997), which leads directly to a second-order differential equation, two-dimensional (Allen 1980, Watts 1993), or three-dimensional (deBoer 1982, Steele 1999, deBoer 2000, Lim and Steele

2002). However, these models could not reproduce all the important aspects of biological cochlea with realistic parameter values. Therefore, other types of cochlear models were proposed. In one type the assumption that the sections of the membrane were coupled only by the fluid was relaxed (Steele 1993, Mammano and Nobili 1993, Geisler and Sang 1995, Nobili and Mammano 1996). In the second type, the membrane was no longer assumed to be moving as a whole, which led to two- and three-mode transmission-line cochlear models (deBoer 1990, Hubbard 1993, Chadwick 1996, Dimitriadis and Chadwick 1999, Hubbard 2000, Lu 2005).

There are numerous analytical methods for solving two- and three-dimensional models. Yet Zweig (Zweig 1991) and Shera (Shera 2005) contend that in the hierarchy of approximations, the cochlear partition representation might be more important than the dimensionality of the longitudinal fluid coupling. The numerical solutions to the multi-mode transmission-line cochlear models are not always straightforward or easy to handle. This hinders the study of the effects of parameter variation or obtaining physical insight.

In section 4.2, we develop an approximate analytical technique for analyzing multi-mode transmission-line cochlear models. We approximate fluid flow to be one-dimensional. This approximation is valid if the wavelength of the traveling wave is large compared to the cross-sectional dimensions of the scalae. Our technique allows to separate the modes and to compute the effective local admittances that produce the corresponding mode in a single-mode one-dimensional transmission-line model. Each mode can then be analyzed separately by any method for single-mode models, including the analytical WKB-approximate solution (Zweig 1991). In section 4.3, we demonstrate the application of our technique to two-mode one-dimensional transmission-line cochlear model (deBoer 1990, Chadwick 1996, Dimitriadis and Chadwick 1999, Hubbard 2000, Lu 2005). We show that the second mode is significant only in the cut-off region, and that the first mode achieves high peaks by having low effective local impedance over an appreciable region basal to and around the CF – a mechanism also observed in other models (Kolston 1990, Zweig 1991, Geisler and Sang 1995, deBoer 2000, Zhak 2004). In section 4.4, we apply our technique to traveling-wave-amplifier cochlear model similar to the one reported in (Hubbard 1993). We show that the second mode is crucial to obtaining high peaks in this model, and it does so by lowering the effective local impedance seen by the first mode over an extended region basal to and around the peak. Our analytical solutions agree qualitatively and quantitatively with the exact numerical simulations, which we use as a standard of comparison for our approximate analytical technique. We conclude and summarize in section 4.5.

4.2. TWO-MODE COCHLEAR MODEL ANALYSIS

Figure 1 shows the general representation of the two-mode one-dimensional transmission-line cochlear model. Voltages P represent pressures, currents U represent volume velocities, and currents I represent linear velocities. The voltages P_1, P_2 and the currents U_1, U_2 satisfy the following two-mode transmission-line equations:

$$-\frac{\partial P_1}{\partial x} = j\omega L_1(x)\cdot U_1 \tag{1a}$$

$$-\frac{\partial P_2}{\partial x} = j\omega L_2(x)\cdot U_2 \tag{1b}$$

$$-\frac{\partial U_1}{\partial x} = I_1 \equiv Y_{11}\cdot P_1 + Y_{12}\cdot P_2 \tag{1c}$$

$$-\frac{\partial U_2}{\partial x} = I_2 \equiv Y_{21}\cdot P_1 + Y_{22}\cdot P_2 \tag{1d}$$

where $L_1(x)$ and $L_2(x)$ are the per-length inductances representing fluid mass in the scalae and $Y_{mn}(j\omega,x)$ are per-length local admittances that depend on specifics of the cochlear model being used. Equations (1a,b) describe macromechanical longitudinal fluid coupling in the cochlea, while (1c,d) represent cochlear model micromechanics.

We assume local scaling symmetry (Zweig 1991), which implies that rather than depending on position and frequency independently, parameters such as local admittances Y_{mn}, voltages and currents in Figure 1 depend only on the following combination of x and ω:

$$\beta(x,\omega) \equiv \frac{\omega}{\omega_c(x)} = \frac{\omega\cdot e^{x/l}}{\omega_c(0)} \qquad s \equiv j\beta \tag{2}$$

where $\omega_c(x)$ is the CF at the location x along the cochlea, and l is the space constant or characteristic length of the exponential cochlear taper; these parameters define the cochlear position-frequency map. Equations (1a-d) then lead to the following coupled ordinary differential equations for P_1 and P_2:

$$\frac{d^2}{ds^2}\begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \begin{bmatrix} a_{11}(s) & a_{12}(s) \\ a_{21}(s) & a_{22}(s) \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \tag{3a}$$

where

$$a_{mn}(s) \equiv Y_{mn}(s) l \cdot \frac{\omega_c(0) \cdot L_m(0) l}{s}, \quad (m,n) \in \{1,2\} \tag{3b}$$

We assume that the matrix elements $a_{11}(s)$, $a_{12}(s)$, $a_{21}(s)$ and $a_{22}(s)$ vary slowly relative to the wavelength of the traveling wave. To solve the Equation (3a) approximately, we consider the solution on a narrow spatial region that corresponds to $s_1<s<s_2$, where $s_1=s_*-\varepsilon$, $s_2=s_*+\varepsilon$, $|\varepsilon|<<|s_*|$. Then the matrix elements can be considered approximately constant over this narrow region:

$$\frac{d^2}{ds^2}\begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \approx \begin{bmatrix} a_{11}(s_*) & a_{12}(s_*) \\ a_{21}(s_*) & a_{22}(s_*) \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \equiv \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \tag{4}$$

The equation (4) is diagonalized by the following linear transform:

$$P_1 = x_1 + b_{12}(s_*) \cdot x_2 \tag{5a}$$

$$P_2 = b_{21}(s_*) \cdot x_1 + x_2 \tag{5b}$$

We set all the diagonal elements of the matrix of the transform (5a,b) to unity for convenience. The equation (4) transforms into:

$$\frac{d^2}{ds^2}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} k_1^2 & 0 \\ 0 & k_2^2 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \tag{6}$$

where $k_1^2(s_*)$ and $k_2^2(s_*)$ are eigenvalues determined from the characteristic polynomial:

$$\left(a_{11} - k_{1,2}^2\right) \cdot \left(a_{22} - k_{1,2}^2\right) - a_{12}a_{21} = 0 \tag{7}$$

The eigenvalues $k_1^2(s_*)$ and $k_2^2(s_*)$ also satisfy the equality that follows from the general properties of characteristic polynomials:

$$k_1^2 + k_2^2 = a_{11} + a_{22} \tag{8}$$

The solution of the quadratic equation in (7) is:

$$k_{1,2}^2\left(s_*\right)=\frac{a_{11}+a_{22}}{2}\pm\sqrt{\left(\frac{a_{11}+a_{22}}{2}\right)^2-\left(a_{11}a_{22}-a_{12}a_{21}\right)} \tag{9}$$

The equation (6) decouples the two modes of wave-propagation, so that each mode can be analyzed separately. The boundary condition that $x_m(s)$ $m\in\{1,2\}$ remain finite as $\beta\rightarrow\infty$ implies that only forward-traveling waves are present (deBoer 1982). We have assumed that the properties of the cochlea scale slowly relative to the wavelength of both traveling waves, so we can use a WKB-type approximation to compute $x_m(s)$:

$$x_m\left(s\right)=c_m\cdot k_m^{-1/2}\left(s\right)\cdot\exp\left(-\int_{s_0}^{s}k_m\left(s'\right)ds'\right) \tag{10}$$

Here $m\in\{1,2\}$; s_0 corresponds to the basal end of the cochlea, and the c_m's are constants that depend on the basal boundary conditions for each mode. To complete our analysis of the problem, we need to compute $b_{12}(s)$ and $b_{21}(s)$ in the equations (5a,b), and then c_1 and c_2 in (10).

We know from linear algebra that the matrix b_{nm} of the transform (5a,b) that diagonalizes the system (4) consists of the eigenvector columns of the matrix $A\equiv\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$ in (4):

$$\left(A-k_m^2\cdot I\right)\cdot b_m=0 \tag{11}$$

where $m\in\{1,2\}$, I is the identity matrix, b_m is the m-th column of the matrix b_{nm} of the transform (5a,b), b_m is also the eigenvector corresponding to the eigenvalue k_m^2.

The system (11) yields the following expressions for $b_{12}(s)$ and $b_{21}(s)$:

$$\begin{aligned} b_{21}\left(s\right)&=\frac{k_1^2-a_{11}}{a_{12}}=\frac{a_{21}}{k_1^2-a_{22}}\\ b_{12}\left(s\right)&=\frac{a_{12}}{k_2^2-a_{11}}=\frac{k_2^2-a_{22}}{a_{21}}\end{aligned} \tag{12}$$

One can imagine the traveling wave in a two-mode cochlear model to be the result of the two modes x_1 and x_2 propagating and slowly rotating along the cochlea to add and form P_1 and P_2 voltages. This slow rotation along the cochlea is defined by the equations (5a,b).

As an example, let us determine the constants c_m for the simple case of a Dirichlet basal boundary condition for P_1. This boundary condition corresponds to having a voltage source V_{in} present at the cochlear input:

$$P_1(s_0) = V_{in} \tag{13}$$

The boundary condition for P_2 is defined by the termination impedance at the basal end of the second line of the cochlea:

$$\alpha_2 \cdot \frac{dP_2}{ds}(s_0) + \beta_2 \cdot P_2(s_0) = 0 \tag{14}$$

Ignoring the pre-exponential dependencies in the equation (10), we can write:

$$\frac{dx_m}{ds}(s) \approx -k_m(s) \cdot x_m(s), \quad m \in \{1,2\} \tag{15}$$

Applying the $\left(\alpha_2 \dfrac{d}{ds} + \beta_2\right)$ operator to the equation (5b) and again assuming that $b_{12}(s)$ and $b_{21}(s)$ vary slowly, the Equations (14) and (15) give us:

$$\frac{x_2(s_0)}{x_1(s_0)} \approx -b_{21}(s_0) \cdot \frac{\beta_2 - \alpha_2 \cdot k_1(s_0)}{\beta_2 - \alpha_2 \cdot k_2(s_0)} \tag{16}$$

Using the boundary condition (13) with the equations (5a), (16) and (10), we calculate the constants c_m:

$$\begin{aligned} \frac{c_2}{c_1} &\approx -b_{21}(s_0) \cdot \frac{k_2^{1/2}(s_0)}{k_1^{1/2}(s_0)} \cdot \frac{\beta_2 - \alpha_2 \cdot k_1(s_0)}{\beta_2 - \alpha_2 \cdot k_2(s_0)} \\ c_1 &\approx V_{in} \cdot k_1^{1/2}(s_0) \Big/ \left(1 - b_{12}(s_0) \cdot b_{21}(s_0) \cdot \frac{\beta_2 - \alpha_2 \cdot k_1(s_0)}{\beta_2 - \alpha_2 \cdot k_2(s_0)}\right) \end{aligned} \tag{17}$$

If $\alpha_2=0$, the boundary condition (14) for P_2 degenerates to a Dirichlet boundary condition $P_2(s_0)=0$, which corresponds to terminating the second line at its basal end with a short circuit.

If $\beta_2=0$, the boundary condition (14) for P_2 degenerates to a Neumann boundary condition $\frac{dP_2}{ds}(s_0) = 0$, which corresponds to terminating the second line at its basal end with an open circuit.

To analyze the two modes in our cochlear model, it is convenient to define the effective local admittances $Y_{eff,m}(s)$, $m \in \{1,2\}$, such that the *single-mode* cochlear model with the characteristic length l, longitudinal fluid coupling $L_1(x)$, and the local admittance $Y_{eff,m}(s)$ would have the wave number $k_m(s)$ defined in the equation (9). Repeating the derivation of the equations (3a,b) for the case of the single transmission line, we obtain the definition of $Y_{eff,m}(s)$:

$$Y_{eff,m}(s) \equiv \frac{s}{\omega_c(0) \cdot L_1(0) \cdot l^2} \cdot k_m^2(s), \quad m \in \{1,2\} \tag{18}$$

The physical meaning of $Y_{eff,m}(s)$ is further exposed by computing the local admittance seen by each line in the two-mode cochlear model. The first line sees the local admittance $Y_1 = I_1/P_1$, and the second line sees $Y_2 = I_2/P_2$. The equations (1c,d), (5a,b) and the definition (3b) then give us:

$$\begin{aligned} Y_1(s) &= \frac{s}{\omega_c(0) \cdot L_1(0) l^2} \cdot \left(a_{11} + a_{12} \cdot \frac{b_{21} \cdot x_1 + x_2}{x_1 + b_{12} \cdot x_2} \right) \\ Y_2(s) &= \frac{s}{\omega_c(0) \cdot L_2(0) l^2} \cdot \left(a_{22} + a_{21} \cdot \frac{x_1 + b_{12} \cdot x_2}{b_{21} \cdot x_1 + x_2} \right) \end{aligned} \tag{19}$$

Often, there are regions of s in the cochlea where one mode dominates. Consider regions where the first mode dominates, i.e., $|x_1(s)| \gg |x_2(s)|$. The equations (19), (12), (8) and the definition (18) give us:

$$\begin{aligned} Y_1(s) &= Y_{eff,1}(s) \\ Y_2(s) &= Y_{eff,1}(s) \cdot \frac{L_1}{L_2} \end{aligned} \tag{20}$$

In the regions where the second mode dominates, i.e., $|x_2(s)| \gg |x_1(s)|$, we similarly get:

$$\begin{aligned} Y_1(s) &= Y_{eff,2}(s) \\ Y_2(s) &= Y_{eff,2}(s) \cdot \frac{L_1}{L_2} \end{aligned} \tag{21}$$

The equations (20) and (21) show that the effective local admittances are not just theoretical variables, but the admittances that each line sees in the regions where the corresponding mode dominates.

Generalization of the presented technique to the case of N-mode one-dimensional transmission-line cochlear models is straightforward. Using Equation (3b) for $(m,n) \in \{1,..N\}$, we obtain a N-by-N matrix instead of the 2-by-2 matrix in Equations (3a) and (4). We diagonalize this N-by-N matrix by solving the characteristic polynomial for eigenvalues k_m^2, $m \in \{1,..N\}$, and using the equation (11) to compute the matrix b_{nm} that consists of N eigenvector columns b_m corresponding to the eigenvalues k_m^2. The matrix b_{nm} transforms the separated modes $x_m(s)$ into the voltages P_n. We calculate these modes $x_m(s)$ in Equation (10). To determine N constants c_m in Equation (10), we impose the boundary condition (13) and $N\text{-}1$ boundary conditions of the form (14). Applying these N basal boundary conditions to the $x_m(s) \rightarrow P_n$ transform and using Equations (15) and (10) at $s=s_0$, we solve for the constants c_m. The definition (18) for $m \in \{1,..N\}$ is still useful, and our result that each line sees a local admittance equal to $Y_{eff,m}(s)$ in the regions where the mode $x_m(s)$ dominates still applies.

4.3. TWO-MODE SANDWICH COCHLEAR MODEL EXAMPLE

4.3.1. Analysis

Figure 2 shows the two-mode one-dimensional Sandwich cochlear model (deBoer 1990, Chadwick 1996, Dimitriadis and Chadwick 1999, Hubbard 2000, Lu 2005). Parts of this model represent physical structures of the biological cochlea such as fluid coupling in the scala vestibuli (SV) and scala tympani (ST), the reticular lamina (RL) and basilar membrane (BM), and outer hair cells (OHCs). Figure 2 shows the electrical circuit representation of the acoustic properties of this cochlear model. In this representation voltages are analogous to acoustic pressures and currents correspond to the velocities. This convention causes parallel mechanical networks to be mapped to series electrical networks and vice versa. In addition, acoustic compliance, viscosity and mass become equivalent to capacitance, resistance and inductance respectively. Capacitances and inductances scale exponentially and resistances stay constant along the length of the cochlea, so that impedances depend only on the combination of x and ω defined in Equation (2). We choose the characteristic frequency $\omega_c(x)$ at the location x along the cochlea to be the local resonant frequency of the BM:

$$\omega_c(x) \equiv \frac{1}{\sqrt{L_{bm}(x) \cdot C_{bm}(x)}} \tag{22}$$

We model V_{active}, the active force generated by the OHCs, as being proportional to the RL deflection, or equivalently to the integrated RL velocity, I_{rl}/s:

$$V_{active} \equiv B_a(s) \cdot \frac{I_{rl}}{s} \qquad (23)$$

where $B_a(s)$ is a proportionality coefficient that depends on s, the combination of x and ω defined in Equation (2). In this model of the biological cochlea, OHCs are assumed to transduce the RL deflection into the potential, which is then low-pass filtered by the OHC membrane. The resultant trans-membrane potential drives OHC force generation. We define the D.C. RL-deflection-to-voltage ratio to be K_v, the D.C. OHC-voltage-to-force gain to be K_f, and the OHC membrane time constant to be T_m. Local scaling symmetry demands that the OHC gain $K_vK_f(x)$ exponentially decrease and the OHC membrane time constant $T_m(x)$ exponentially increase with x such that B_a depends only on s (deBoer 1990, Chadwick 1996, Lu 2005):

$$B_a \equiv \frac{B_t}{1+s\tau_m} \qquad (24)$$

The definitions and values of the dimensionless parameters that we use for this cochlear model are given in Table 1. The parameter values are similar to those used in (Lu 2005) and measured in the papers cited therein. Note that the technique that we developed in Section II is general; it works for any model described by Figure 1, and for any parameter values. The values in Table 1, therefore, are for illustrative purposes only.

We define the following normalized functions to be impedances of the BM, RL and OHC respectively multiplied by $s/\left(\omega_c(0)L_{bm}(0)\right)$ (see Figure 2):

$$Z_{bm} \equiv s^2 + s/Q_{bm} + 1 \qquad (25)$$

$$Z_{rl} \equiv M \cdot \left(s^2 + s \cdot \omega_{rl}/Q_{rl} + \omega_{rl}^2\right) \qquad (26)$$

$$Z_{ohc} \equiv s \cdot c \cdot M \cdot \omega_{rl}/Q_{rl} + K \qquad (27)$$

We can derive the expressions for I_{rl} and I_{bm}, given the voltages P_1 and P_2:

$$I_{rl} = \frac{s}{\omega_c(0)L_{bm}(0) \cdot Z_Z(s)} \cdot \left[\left(Z_{bm} + Z_{ohc}\right) \cdot P_1 - Z_{ohc} \cdot P_2\right] \qquad (28)$$

$$I_{bm} = \frac{s}{\omega_c(0) L_{bm}(0) \cdot Z_Z(s)} \cdot \left[-\left(Z_{ohc} + B_a\right) \cdot P_1 + \left(Z_{rl} + Z_{ohc} + B_a\right) \cdot P_2 \right] \tag{29}$$

where

$$Z_Z(s) \equiv Z_{rl} \cdot \left(Z_{bm} + Z_{ohc}\right) + Z_{bm} \cdot \left(Z_{ohc} + B_a\right) \tag{30}$$

Equations (28) and (29) have the same form as equations (1c, d), so we can relate the admittances $Y_{mn}(s)$ to parameters of this model. Equation (3b) then yields:

$$\frac{d^2}{ds^2} \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} = \frac{(4N)^2}{Z_Z(s)} \cdot \begin{bmatrix} Z_{bm} + Z_{ohc} & -Z_{ohc} \\ -R \cdot \left(Z_{ohc} + B_a\right) & R \cdot \left(Z_{rl} + Z_{ohc} + B_a\right) \end{bmatrix} \cdot \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \tag{31}$$

Equation (31) is of the same form as (3a), so we can apply the technique developed in Section II.

The second line is terminated at the basal end with an open circuit. This termination corresponds to $\beta_2=0$ in the boundary condition (14). We now substitute $\beta_2=0$ into Equation (17) to compute the constants c_m.

Bounded-input bounded-output (BIBO) stability of this cochlear model was checked for the parameter values in Table 1. A standard state-space system representation can be used to investigate BIBO stability because the model comprises a finite number of elements with rational frequency-domain impedances. The necessary and sufficient condition for BIBO stability of the system is that all the eigenvalues of its state-space matrix have negative real parts. The active gain parameter B_t was tapered down towards the apex to reduce apical reflections and improve stability at low frequencies.

4.3.2. Results

Equations (9), (31) and $Y_{eff,m}(s)=1/Z_{eff,m}(s)$ (from Equation (18), with L_{sv} replacing L_l) yield the following expression for the effective local impedances normalized by $\omega_c(0)L_{bm}(0)$:

$$sZ_{eff1,2}(s) = \frac{Z_Z(s) \cdot \left(1 + s\tau_m\right)}{Z_1(s) \pm \sqrt{Z_1^2(s) - R \cdot \left[Z_Z(s) \cdot \left(1 + s\tau_m\right)\right] \cdot \left(1 + s\tau_m\right)}} \tag{32}$$

where:

$$Z_1(s) \equiv \left[Z_{bm} + Z_{ohc} + R \cdot \left(Z_{rl} + Z_{ohc} + B_a\right)\right] \cdot \left(1 + s\tau_m\right)/2 \tag{33}$$

To visualize Equations (9) and (32), we would like to approximate $sZ_{eff1,2}(s)$ in Equation (32) by using rational functions. The numerator in Equation (32), i.e., $Z_Z(s)\cdot(1+s\tau_m)$, is a fifth-order polynomial. Therefore, $sZ_{eff1,2}(s)$ has five zeros. The function $Z_I(s)$ defined by the Equation (33) is a third-order polynomial, making the expression under the square root in Equation (32) a sixth-order polynomial. Therefore, the square root in the Equation (32) behaves like a third-order polynomial at very low and very high frequencies ($s\to j0$ and $s\to j\infty$). So we attempt to approximate the denominator of the Equation (32) using a third-order polynomial of the form $p_1\cdot(s+p_2)\cdot(s^2+s\cdot p_3+p_4)$. This simple approximation works very well for the first mode: both real and imaginary parts of $Z_{eff1}(s)$ (from Equation (32)) match their rational approximations closely over a wide range of frequencies. The pole-zero plot of this rational approximation to $sZ_{eff1}(s)$ is shown in Figure 3 (A). We observe two zero pairs and a pole pair close to the imaginary axis. This structure is similar to that seen in (Zweig 1991, Zhak 2004).

Approximating the denominator of Equation (32) by a similarly simple third-order polynomial does not work for the second mode. The four degrees of freedom that are offered by the four coefficients in the third-order polynomial allow us to match the real and imaginary parts of $Z_{eff2}(s)$ for the very high and very low frequencies s, but the match around $s=j1$ is inadequate. To increase the number of degrees of freedom, we use a Pade-like rational approximation to the denominator of the Equation (32) for the second mode: $\dfrac{r_1\cdot(s+r_2)\cdot(s^2+s\cdot r_3+r_4)\cdot(s^2+s\cdot r_5+r_6)}{s^2+s\cdot r_7+r_8}$.

Now we are able to match real and imaginary parts of $Z_{eff2}(s)$ for a wide range of frequencies s. The pole-zero plot of this rational approximation to $sZ_{eff2}(s)$ is shown in Figure 3 (B). We see that the poles due to $r_1\cdot(s+r_2)\cdot(s^2+s\cdot r_3+r_4)\cdot(s^2+s\cdot r_5+r_6)$ in the denominator cancel out the zeros due to the numerator polynomial $Z_Z(s)\cdot(1+s\tau_m)$ almost exactly. Therefore, we are left with just a zero pair for $sZ_{eff2}(s)$. The rational approximation $sZ_{eff2}\approx q_1\cdot(s^2+s\cdot q_2/q_3+q_2^2)$ provides an excellent match to the sZ_{eff2} computed in Equation (32) for the second mode over a wide frequency range. The pole-zero plot of this very simple rational approximation to $sZ_{eff2}(s)$ is shown in Figure 3 (C). The zero pair corresponds to propagation in a single-mode cochlear model with local impedance resulting from a simple resonator with relatively low Q. Therefore, we expect the first mode to be primarily responsible for cochlear amplification near the peak of the amplitude response.

Figure 4 shows the modal decomposition of the RL shunt current I_{rl}. The bold solid line shows I_{rl} amplitude when the two-mode Sandwich cochlear model described by Equation (31) is solved exactly. The dashed line shows I_{rl} amplitude calculated using the technique described in Section II assuming only the first mode is present. The solid line shows I_{rl} amplitude when only the second mode is present. Note that the second mode, which is due to longitudinal fluid coupling $L_{st}(x)$ in the scala tympani, contributes significantly to the solution only in the cut-off region. It has negligible effect on cochlear amplification near the peak, as expected from the pole-zero plots in Figure 3 (A) and (C).

Figure 5 demonstrates good agreement between the exact (solid line) solution for current I_{rl} and the approximate (dashed line) solution computed using the technique described in Section II. Figure 5 (A) shows the amplitude of the current I_{rl}, and Figure 5 (B) shows its phase.

Figure 6 shows the modal decomposition of the BM shunt current I_{bm}. The bold solid line shows the I_{bm} amplitude when Equation (31) is solved exactly. The dashed line shows I_{bm} amplitude calculated using the technique described in Section II when only the first mode is present. The solid line shows I_{bm} amplitude when only the second mode is present. Again, the second mode has negligible effect on cochlear amplification near the peak, contributing to the solution only in the cut-off region.

Figure 7 further illustrates good agreement between the exact (solid line) solution for current I_{bm} and the approximate (dashed line) solution. Figure 7 (A) shows the amplitude of the current I_{bm}, and Figure 7 (B) shows its phase.

Figure 8 shows the local impedance P_1/I_{rl} seen by the first line in the two-mode cochlear model, demonstrating good agreement between the exact solution (solid line) of Equation (31), and the approximate solution (dashed line) described in Section II. The top panel shows the real part of the impedance (resistance), and the bottom panel shows the imaginary part (reactance). The first mode dominates at every s except in the cut-off region, so the approximate impedance in Figure 8 is very close to $1/Y_{eff1}(s)$. The rough, but rather accurate approximation to $k_1^2(s)$ in Sandwich cochlear model is $a_{11}(s)$, because in this model the coupling between the two modes is weak. This approximation corresponds to shorting P_2 to ground. Thus, cochlear amplification can be analyzed by considering a single-mode cochlear model with a local impedance $Z_1(s)$ given by

$$\frac{s}{\omega_c(0)L_{bm}(0)}\cdot Z_1(s)\approx Z_{rl}+\frac{Z_{bm}\cdot Z_{ohc}}{Z_{bm}+Z_{ohc}}+\frac{B_a\cdot Z_{bm}}{Z_{bm}+Z_{ohc}} \qquad (34)$$

The EQ-NL theorem (deBoer 1997, deBoer 2000) can be applied to the approximation (34) to gain intuition about the cochlear amplifier in the Sandwich model. We interpret Figure 8 as showing the local impedance along the length of the cochlea for fixed frequency. The effect of the cochlear amplifier is to reduce the magnitude of the local impedance over an extended region basal to and around the peak, allowing high peaks in I_{rl} to be achieved without significantly increasing voltage P_1 and infusing excessive amounts of energy into the traveling wave.

The basal termination of the second line does not affect peak gain in this cochlear model, because the output currents I_{rl} and I_{bm} are determined by the first mode everywhere except in the cut-off region.

4.4. TRAVELING-WAVE-AMPLIFIER COCHLEAR MODEL EXAMPLE

4.4.1. Analysis

Figure 9 shows the two-mode one-dimensional traveling-wave-amplifier cochlear model similar to the one reported in (Hubbard 1993). The only difference between the model investigated in this Section and the one reported in (Hubbard 1993) is the sign of the feedback between the two transmission lines. It turns out that the model can have cochlea-like responses and be BIBO stable for both signs of the feedback. However, choosing the values of the parameters of the model to obtain frequency responses reported in (Hubbard 1993) leads to singularities of the matrix A in (4) near jω-axis. Our assumptions about the matrix A and the WKB approximation break down near the corresponding frequencies, which might indicate reflections of various modes invalidating the reasoning that lead to (10). Therefore, we only consider one model in this Section, and the work to incorporate reflections and other interactions among the modes near the regions where our assumptions about the matrix A and the WKB approximation break down is still ongoing.

In the two-mode one-dimensional traveling-wave-amplifier cochlear model shown in Figure 9 the capacitances and inductances scale exponentially and resistances stay constant along the length of the cochlea, so that impedances depend only on β, the combination of x and ω defined in (2). We choose the characteristic frequency $\omega_c(x)$ at location x along the cochlea to be the local resonant frequency of the first line:

$$\omega_c(x) \equiv \frac{1}{\sqrt{L_0(x)\cdot C_0(x)}} \tag{35}$$

The definitions and values of the dimensionless parameters that we use for this cochlear model are given in Table 2. These values are only illustrative; the generic technique described in Section II works for any model represented by Figure 1 and for arbitrary parameter values, as long as the assumptions about the matrix $\boldsymbol{A}$ in (4) hold.

Physically, parameter D represents the ratio of the group velocities (at low frequencies) of the first (resonant) and second transmission lines. Parameter γ_a represents coupling between the lines.

Deriving the expressions for $I_1 = I_{out}$ and I_2 as functions of P_1 and P_2 to compute local admittances $Y_{mn}(s)$, and applying (3b), we obtain:

$$\frac{d^2}{ds^2}\begin{bmatrix} P_1 \\ V_2 \end{bmatrix} = \frac{(4N)^2}{s\cdot\left(s^2+s/Q_1+1\right)}\cdot\begin{bmatrix} s & \gamma_a\cdot s \\ -D^2 & D^2\cdot\left\{\left(s+\omega_2\right)\cdot\left(s^2+s/Q_1+1\right)-\gamma_a\right\} \end{bmatrix}\cdot\begin{bmatrix} P_1 \\ V_2 \end{bmatrix} \tag{36}$$

where:

$$V_2 \equiv P_2\cdot\frac{\omega_c(0)\cdot C_2(0)}{G_{12}} \tag{37}$$

The output current I_{out} is given by:

$$I_{out} = \frac{s}{\omega_c(0)\cdot L_0(0)}\cdot\frac{P_1+\gamma_a\cdot V_2}{s^2+s/Q_1+1} \tag{38}$$

Equation (36) is of the same form as (3a), so we apply the technique developed in Section II.

The second line is terminated at the basal end with a resistance $Z_t = \sqrt{L_2(0)/C_2(0)}$ that is approximately equal to its characteristic impedance. Substituting (1b) into the boundary condition (14) and noting that $P_2(s_0)/U_2(s_0) = -Z_t$, we obtain:

$$\beta_2/\alpha_2 = -4N \cdot D \tag{39}$$

We can now substitute (39) into (17) to compute the constants c_m.

Bounded-input bounded-output (BIBO) stability of this cochlear model was checked for the parameter values listed in Table 2. We can use a standard state-space system representation for evaluating stability because the model comprises a finite number of lumped elements. We gradually introduced viscosity in series with $L_1(x)$ and $L_2(x)$ near the apical end to reduce the amplification of the cochlear sections (Xin 2003). This gain tapering reduces apical reflections and therefore improves the stability of our cochlear model at low frequencies.

4.4.2. Results

Equations (9), (36) and the definition (18) for $Y_{eff,m}(s)=1/Z_{eff,m}(s)$ yield the following expression for effective local impedances normalized by $\omega_c(0)L_0(0)$:

$$sZ_{eff1,2}(s) = \frac{s \cdot \left(s^2 + s/Q_1 + 1\right)}{Z_1(s) \pm \sqrt{Z_1^2(s) - D^2 \cdot s \cdot \left(s^2 + s/Q_1 + 1\right) \cdot \left(s + \omega_2\right)}} \tag{40}$$

where:

$$Z_1(s) \equiv \left[s + D^2 \cdot \left\{ \left(s + \omega_2\right) \cdot \left(s^2 + s/Q_1 + 1\right) - \gamma_a \right\} \right] / 2 \tag{41}$$

As before, we approximate $sZ_{eff1,2}(s)$ using rational functions. The function $Z_1(s)$ defined by (41) is a third-order polynomial, so the expression under the square root in (40) is a sixth-order polynomial with coefficients of s^6 and s^5 equal to those of $Z_1^2(s)$. Therefore, the square root in (40) behaves at high frequencies like a third-order polynomial with coefficients of s^3 and s^2 equal to those of $Z_1(s)$. For the first mode the terms with s^3 and s^2 in the denominator of (40) will therefore cancel out. However, approximating the denominator with a first-order polynomial does not offer enough degrees of freedom to match both real and imaginary parts of $Z_{eff1}(s)$ for a wide range of frequencies. To increase the number of degrees of freedom, we use a Pade-like rational approximation to the denominator of (40) for the first mode, as follows: $\dfrac{r_1 \cdot \left(s + r_2\right) \cdot \left(s^2 + s \cdot r_3 + r_4\right)}{s^2 + s \cdot r_5 + r_6}$. The pole-zero plot of this rational approximation to $sZ_{eff1}(s)$ is shown in Figure 10 (A). We observe that the pole pair due to

$s^2+s\cdot r_3+r_4$ in the denominator cancels out the zero pair due to s^2+s/Q_1+1 in the numerator almost exactly. So the quality factor, or sharpness, of the first mode is much lower than Q_1.

For the second mode approximating the denominator of the (40) by a third-order polynomial of the form $p_1\cdot(s+p_2)\cdot(s^2+s\cdot p_3+p_4)$ works very well. Both real and imaginary parts of $Z_{eff2}(s)$ (from (40)) closely match its rational approximation over a wide frequency range. The pole-zero plot of this rational approximation to $sZ_{eff2}(s)$ is shown in Figure 10 (B). Note that the square root in (40) behaves at high frequencies like a third-order polynomial with coefficients of s^3 and s^2 equal to those of $Z_1(s)$ in (41). It makes the factor p_1 close to D^2, reducing the effective local impedance $Z_{eff2}(s)$ by the same factor and increasing the effective wave number $k_2(s)$ of the second mode by the factor D. This factor D appears under the exponential in (10) and affects the gain and phase shift of the second mode. In this cochlear model, we therefore expect the second mode to be primarily responsible for amplification near the peak.

Figure 11 shows the modal decomposition of current I_{out}. The bold solid line shows I_{out} amplitude when (36) is solved exactly. The dashed line shows I_{out} amplitude calculated using the technique described in Section II when only the first mode is present. The solid line shows I_{out} amplitude when only the second mode is present. Interference effects caused by interaction between the first and the second modes can be seen at $\beta<0.7$, although the first mode is dominant in this region. The second mode dominates in the region $0.7<\beta<1.1$ and thus determines the active gain of this cochlear model. The first mode dominates again in the cut-off region for $\beta>1.1$.

Figure 12 demonstrates excellent agreement between the exact (solid line) solution for current I_{out} and the approximate (dashed line) solution computed using the technique described in Section II. The top panel shows the amplitude of the current I_{out}, and the bottom panel shows its phase.

Figure 13 shows the modal decomposition of local impedance P_1/I_{out} seen by the first line in this two-mode cochlear model. Figure 13 (A) shows the real part of the impedance (resistance), and Figure 13 (B) shows the imaginary part (reactance). The solid line in Figure 13 (A) and (B) shows the local impedance computed from the exact solution, while the dashed line shows the effective local impedance of the first mode, $Z_{eff1}(s)=1/Y_{eff1}(s)$. The bold solid line shows the effective local impedance of the second mode, $Z_{eff2}(s)=1/Y_{eff2}(s)$. In the region $\beta>1.1$ the first mode dominates, so the exact impedance follows the effective impedance of the first mode. In the region $\beta<0.7$ we notice interference between the first and the second modes, with the first mode dominant. In the region $0.7<\beta<1.1$ the second mode dominates, and the exact impedance follows the effective impedance of the second mode. The real and imaginary parts of this effective impedance are shown with greater resolution in Figure 13 (C). The resistance is negative over the region $0.7<\beta<0.95$, basal to the peak, which corresponds to energy transfer into the traveling wave. As expected, the magnitude of the effective local impedance of the second mode is greatly reduced because of the $1/D^2$ factor. This reduction allows high peaks in I_{out} to be achieved without significantly increasing voltage P_1 and infusing excessive amounts of energy into the traveling wave from the first line. Another way of explaining high peaks in I_{out} was discussed in (Hubbard 1993). That reasoning still holds for the traveling-wave-amplifier that we investigate in this Section that has the opposite sign of the feedback between the two transmission lines. The traveling wave in the first line slows down around its resonant location, making its group velocity closer to that in the second line, causing coherent excitation of the second line. This coherent excitation causes the amplitude of P_2 to rise sharply. It can be seen as follows: on Thevenizing the transconductor G_{21} and resistor R_1, we obtain a voltage source proportional to P_2 that drives the current I_{out} even if the amplitude of P_1 does not increase.

Figure 14 demonstrates excellent agreement between the local impedance P_1/I_{out} computed from the exact solution (solid line) of (36), and the approximate solution (dashed line) found using our technique described in Section II. The top panel shows the real part of the local impedance (resistance), and the bottom panel shows the imaginary part (reactance).

In this cochlear model the value of c_2 is very sensitive to β_2/α_2, because $\left|k_1(s_0)/k_2(s_0)\right| \square\ 1$. Therefore, the termination of the second line at its basal end is very important for determining the peak gain of this cochlear model. For example, terminating the second line with an open circuit would significantly degrade the peak gain.

4.5. CONCLUSIONS

We have developed an approximate analytical technique for analyzing multi-mode one-dimensional transmission-line cochlear models. This technique allows separating the modes. For each mode, we have computed the effective local admittance that would produce that mode in a single-mode one-dimensional transmission-line model, which can then be analyzed by any method for single-mode models, including the analytical WKB-approximate technique. We have demonstrated the application of our technique to two-mode Sandwich cochlear model, obtaining an important physical insight that the second mode is significant only in the cut-off region. We have also applied our technique to traveling-wave-amplifier cochlear model, showing that the second transmission line is crucial to achieving high peaks, doing so by lowering the effective local impedance seen by the first line over an appreciable region basal to and around the peak. Our analytical solutions agree qualitatively and quantitatively with the exact numerical simulations, which we use as a standard of comparison for our approximate technique.

4.6. REFERENCES

Allen, J. B. **(1980)** "Cochlear micromechanics – a physical model of transduction," *J. Acoust. Soc. Am.* **68**:1660-1670.

Chadwick, R. S., Dimitriadis, E. K. and Iwasa E. K. **(1996)** "Active control of waves in a cochlear model with subpartitions," *Proc. Natl. Acad. Sci. U.S.A.* **93**:2564-2569.

de Boer, E., and van Bienema, E. **(1982)** "Solving cochlear mechanics problems with higher-order differential equations," *J. Acoust. Soc. Am.* **72**:1427-1434.

de Boer, E. **(1990)** "Wave-propagation modes and boundary conditions for the Ulfendahl-Flock-Khanna (UFK) preparation." *In Mechanics and Biophysics of Hearing*. P. Dallos, C. D. Geisler, J. W. Matthews, M. A. Ruggero, and C.R. Steele, editors. Spring-Verlag, Berlin, pp. 333-339.

de Boer, E. **(1997)** "Connecting frequency selectivity and nonlinearity for models of the cochlea," Aud. Neurosci. **3**:377-388.

de Boer, E., and A. L. Nuttall **(2000)** "The mechanical waveform of the basilar membrane. III. Intensity effects," *J. Acoust. Soc. Am.* **107**:1497-1507.

Dimitriadis, E. K., Chadwick, R. S. **(1999)** "Solution of the inverse problem for a linear cochlear model: A tonotopic cochlear amplifier," *J. Acoust. Soc. Am.* **106:**1880-1892.

Fukazawa, T. **(1997)** "A model of cochlear micromechanics," *Hear. Res.* **113**:182-190.

Geisler, C. D. and C. Sang. **(1995)** "A cochlear model using feed-forward outer hair cell forces," *Hear. Res.* **86**:132-146.

Hubbard, A. E. **(1993)** "A traveling-wave amplifier model of the cochlea," *Science* **259:**68-71

Hubbard, A. E., Yang, Z., Shatz, L., and Mountain, D. C. **(2000)** "Multi-mode cochlear models," *In* Symposium on Recent Developments in Auditory Mechanics. H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, and T. Koike, editors. World Scientific Publishing, Singapore, pp. 167-173.

Kolston, P. J., de Boer, E., Viergever, M. A., Smoorenburg, G. F. **(1990)** "What type of force does the cochlear amplifier produce?" *J. Acoust. Soc. Am.* **88**:1794-1801.

Lim, K. M., and C. R. Steele **(2002)** "A three-dimensional nonlinear active cochlear model analyzed by the WKB-numeric method," *Hear. Res.* **170**:190-205.

Lu, T. K., Zhak, S., Dallos, P., and R. Sarpeshkar **(2005)** "A micromechanical model for fast cochlear amplification with slow outer hair cells," *In* Auditory Mechanisms: Processes and Models. World Scientific Publishing, Singapore. In press.

Mammano, F., Nobili, R. **(1993)** "Biophysics of the cochlea: Linear approximation," *J. Acoust. Soc. Am.* **93**:3320-3332.

Neely, S. T. **(1993)** "A model of cochlear mechanics with outer hair cell motility," *J. Acoust. Soc. Am.* **94**:137-146.

Nobili, R., and F. Mammano **(1996)** "Biophysics of the cochlea. II: Stationary nonlinear phenomenology," *J. Acoust. Soc. Am.* **99**:2244-2255.

Ruggero, M., Rich, N., Robles, L., and B. Shivapuja **(1990)** "Middle-ear response in the chinchilla and its relationship to mechanics at the base of the cochlea," *J. Acoust. Soc. Am.* **87**:1612–1629.

Shera, C. A., Tubis, A., Talmadge, C. L. (**2005**) "Coherent reflection in a two-dimensional cochlea: Short-wave versus long-wave scattering in the generation of reflection-source otoacoustic emissions," *J. Acoust. Soc. Am.* **118**:287-313.

Steele, C. F., Baker, G., Tolomeo, J. and Zetes, D. (**1993**) "Electromechanical models of the outer hair cell," In Biophysics of Hair Cell Sensory Systems. H. Duifhuis, J. W. Horst, P. van Dijk and S. M. van Netten (Eds.), World Scientific, pp. 207-215.

Steele, C. R. (**1999**) "Toward three-dimensional analysis of cochlear structure," *Otorhinolaryngol. Relat. Spec.,* **61**: 238-251.

Watts, L. (**1993**) "Cochlear Mechanics: Analysis and Analog VLSI," *Ph.D. Thesis*, Caltech.

Xin, J., Qi, Y-Y., and Deng, L. (**2003**) "Time domain computation of a nonlinear nonlocal cochlear model with applications to multitone interaction in hearing," Comm. Math. Sci., **1**(2):211-227.

Zhak, S. M., Mandal, S., and Sarpeshkar, R. (**2004**) "A proposal for an RF cochlea," *in Proc. Asia Pacific Microwave Conference.*

Zweig, G. (**1991**) "Finding the impedance of the organ of Corti," *J. Acoust. Soc. Am.*, **89:**1229-1254.

TABLE 1. Parameter definitions and values that we use for the two-mode Sandwich cochlear model.

Parameter	Definition	Value
N	$(4N)^2 \equiv \dfrac{L_{sv}(x)\cdot l}{L_{bm}(x)/l}$	1.5
R	$R \equiv \dfrac{L_{st}(x)}{L_{sv}(x)}$	1.5
B_t	$B_t \equiv K_f K_v(x)\cdot C_{bm}(x)$	0.7
τ_m	$\tau_m \equiv \omega_c(x)\cdot T_m(x)$	33
ω_{rl}	$\omega_{rl} \equiv \dfrac{\sqrt{L_{bm}(x)C_{bm}(x)}}{\sqrt{L_{rl}(x)C_{rl}(x)}}$	0.76
Q_{rl}	$Q_{rl} \equiv \dfrac{1}{R_{rl}}\cdot\sqrt{\dfrac{L_{rl}(x)}{C_{rl}(x)}}$	3.8
Q_{bm}	$Q_{bm} \equiv \dfrac{1}{R_{bm}}\cdot\sqrt{\dfrac{L_{bm}(x)}{C_{bm}(x)}}$	4.4
M	$M \equiv \dfrac{L_{rl}(x)}{L_{bm}(x)}$	0.08
K	$K \equiv \dfrac{C_{bm}(x)}{C_{ohc}(x)}$	0.04
c	$c \equiv \dfrac{R_{ohc}}{R_{rl}}$	0

TABLE 2. Parameter definitions and values that we use for the two-mode traveling-wave-amplifier cochlear model.

Parameter	Definition	Value
N	$(4N)^2 \equiv \dfrac{L_1(x)\cdot l}{L_0(x)/l}$	0.75
D	$D \equiv \dfrac{\sqrt{L_2(x)C_2(x)}}{\sqrt{L_1(x)C_0(x)}}$	50
γ_a	$\gamma_a \equiv \dfrac{G_{12}G_{21}\cdot R_1}{\omega_c(x)\cdot C_2(x)}$	0.28
Q_1	$Q_1 \equiv \dfrac{1}{R_1}\cdot\sqrt{\dfrac{L_0(x)}{C_0(x)}}$	4
ω_2	$\omega_2 \equiv \dfrac{1}{\omega_c(x)\cdot C_2(x)\cdot R_2}$	0.33

FIG. 1. The generic two-mode one-dimensional transmission-line cochlear model.

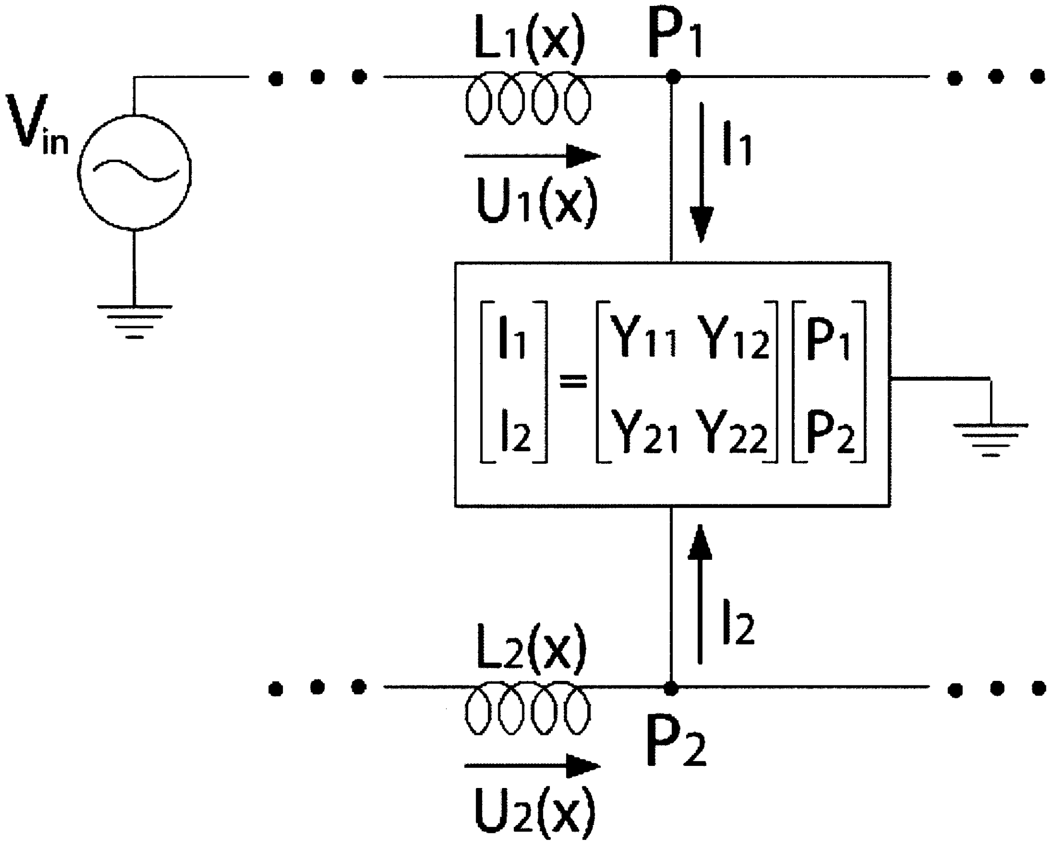


FIG. 2. The two-mode Sandwich cochlear model (Chadwick 1996, Dimitriadis and Chadwick 1999, Hubbard 2000, Lu 2005).

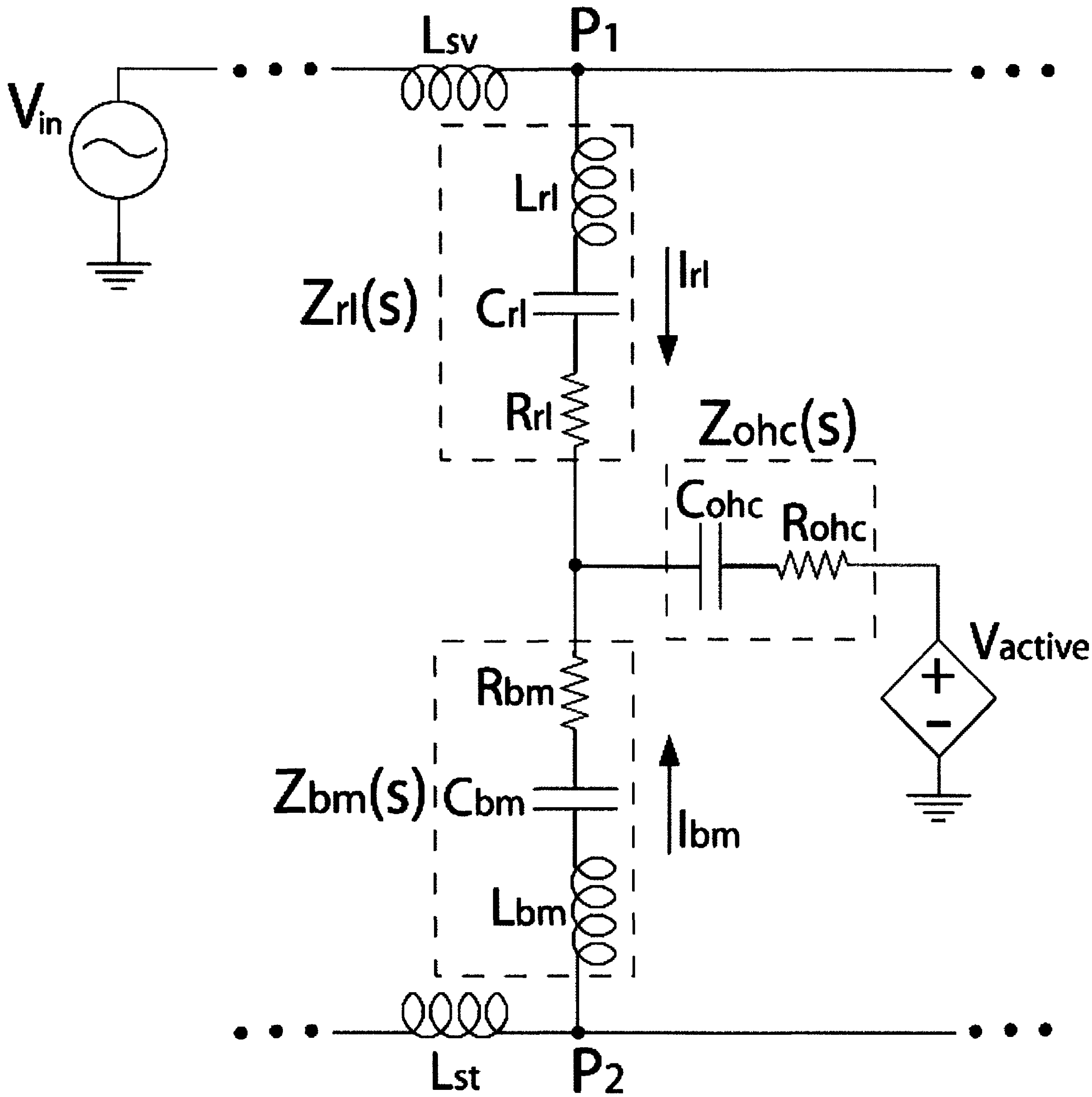


FIG. 3. The pole-zero plot of the rational approximation to $sZ_{eff1,2}(s)$ in the two-mode Sandwich cochlear model: (A) The first (dominant) mode, (B) the second mode, (C) the second mode, simplified.

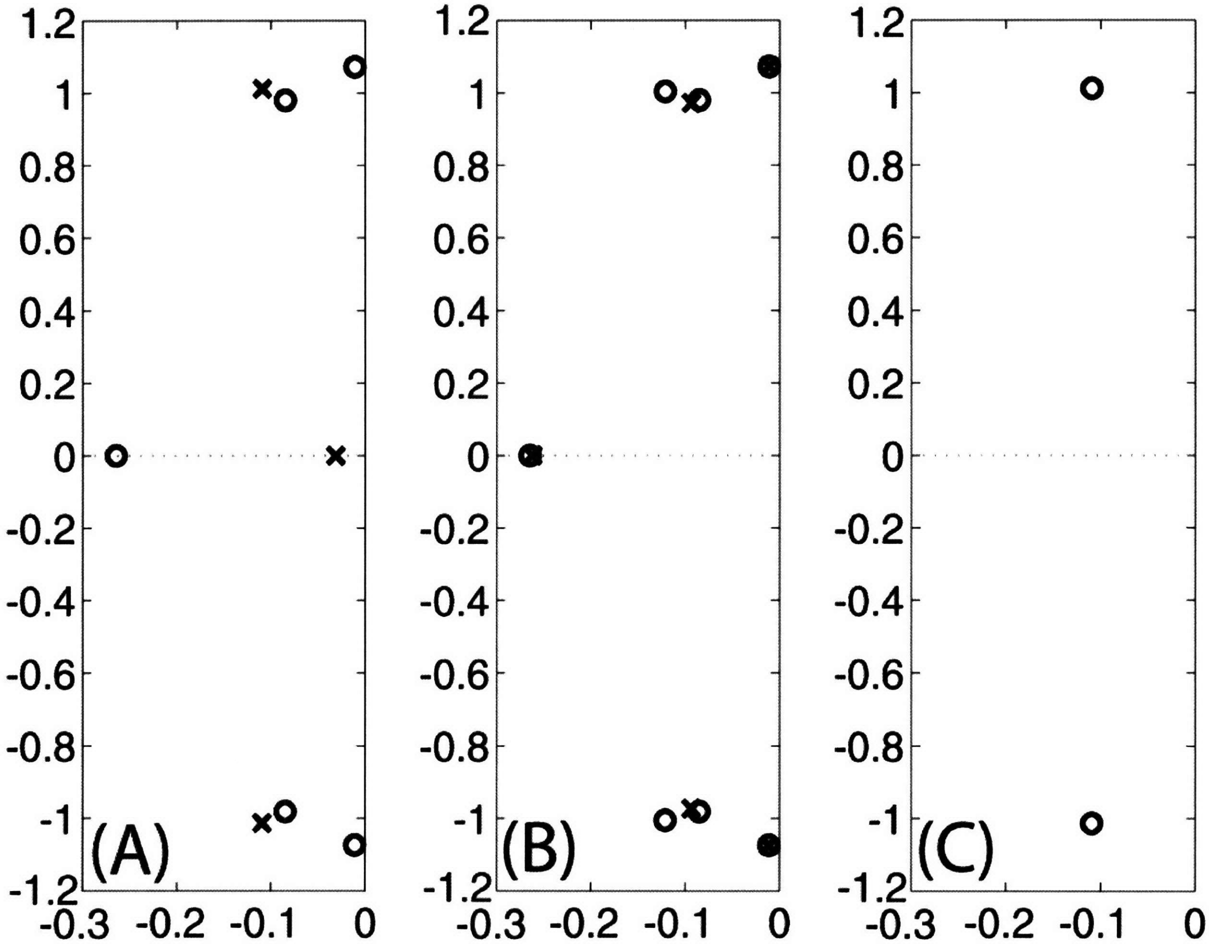


FIG. 4. Modal decomposition of I_{rl} amplitude in the two-mode Sandwich model: (Bold solid) shows the exact solution; (Dashed) shows the first mode; (Solid) shows the second mode.

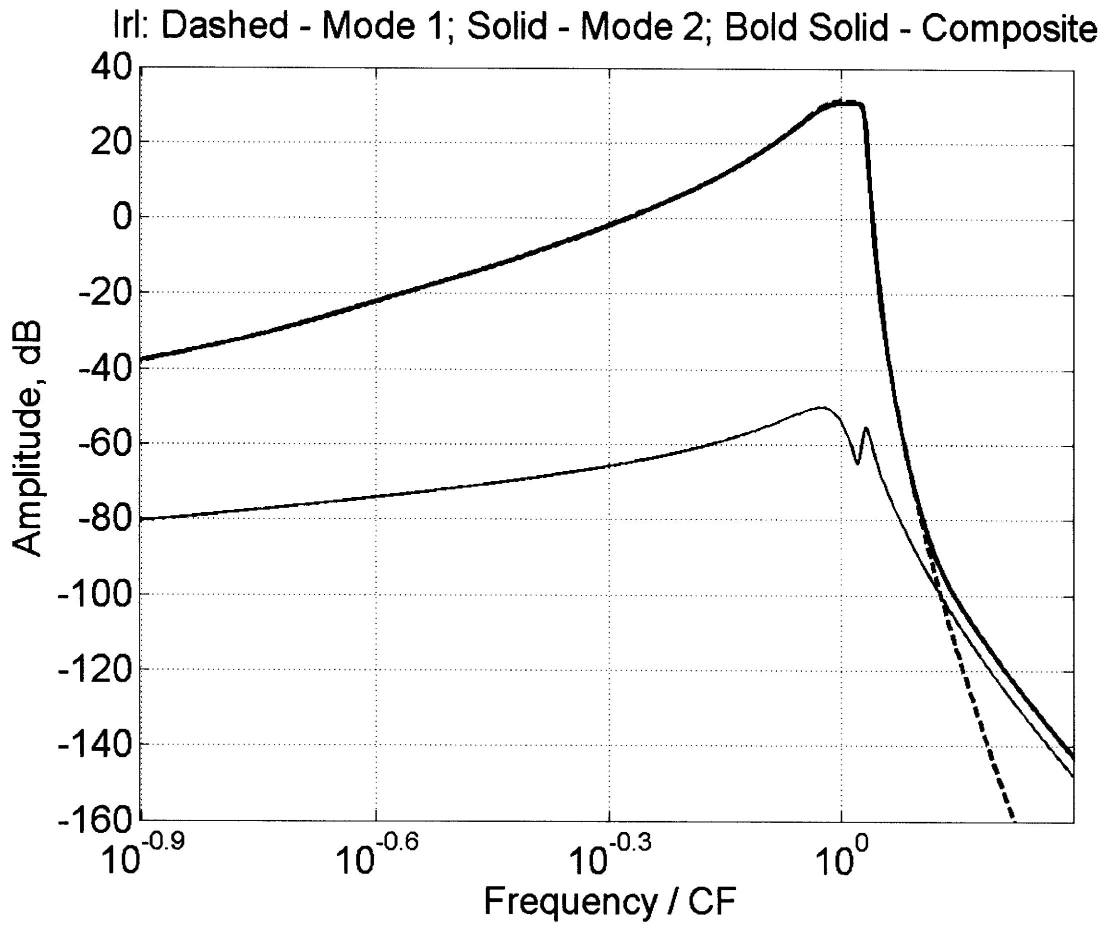


FIG. 5. I_{rl} in the two-mode Sandwich model: (Solid) shows the exact solution, (Dashed) shows the approximation described in Section II; (A) Amplitude (dB), and (B) Phase (cycles).

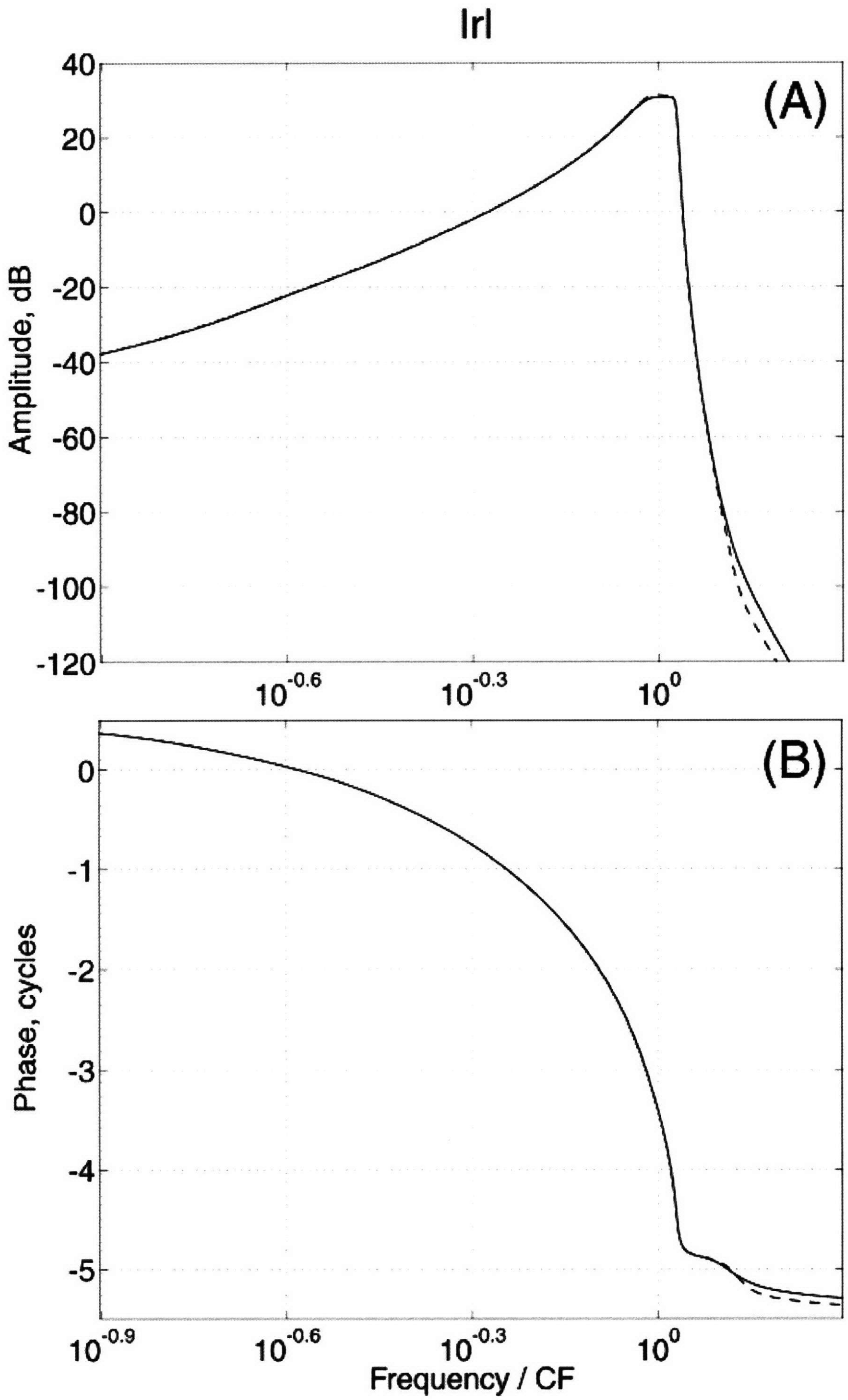


FIG. 6. Modal decomposition of I_{bm} amplitude in the two-mode Sandwich model: (Bold solid) shows the exact solution; (Dashed) shows the first mode; (Solid) shows the second mode.

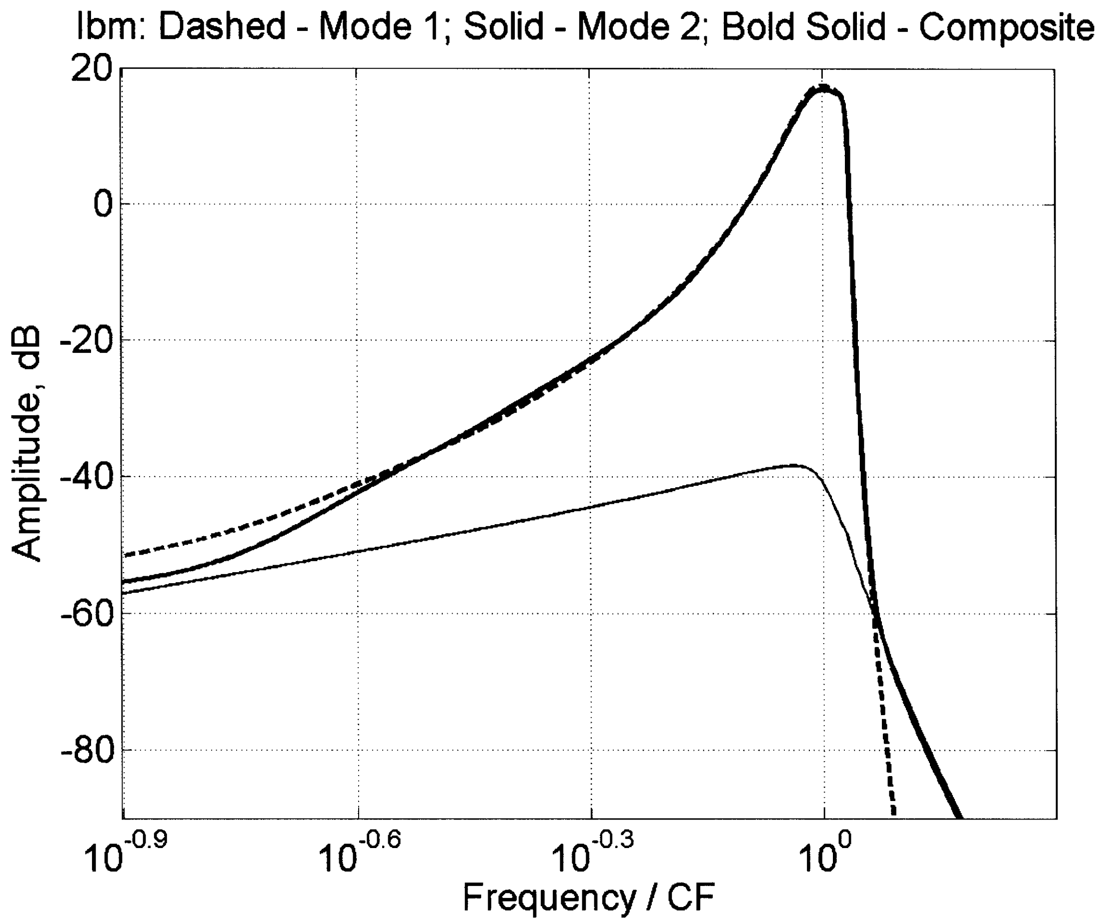


FIG. 7. I_{bm} in the two-mode Sandwich model: (Solid) shows the exact solution, (Dashed) shows the approximation described in Section II; (A) Amplitude (dB), and (B) Phase (cycles).

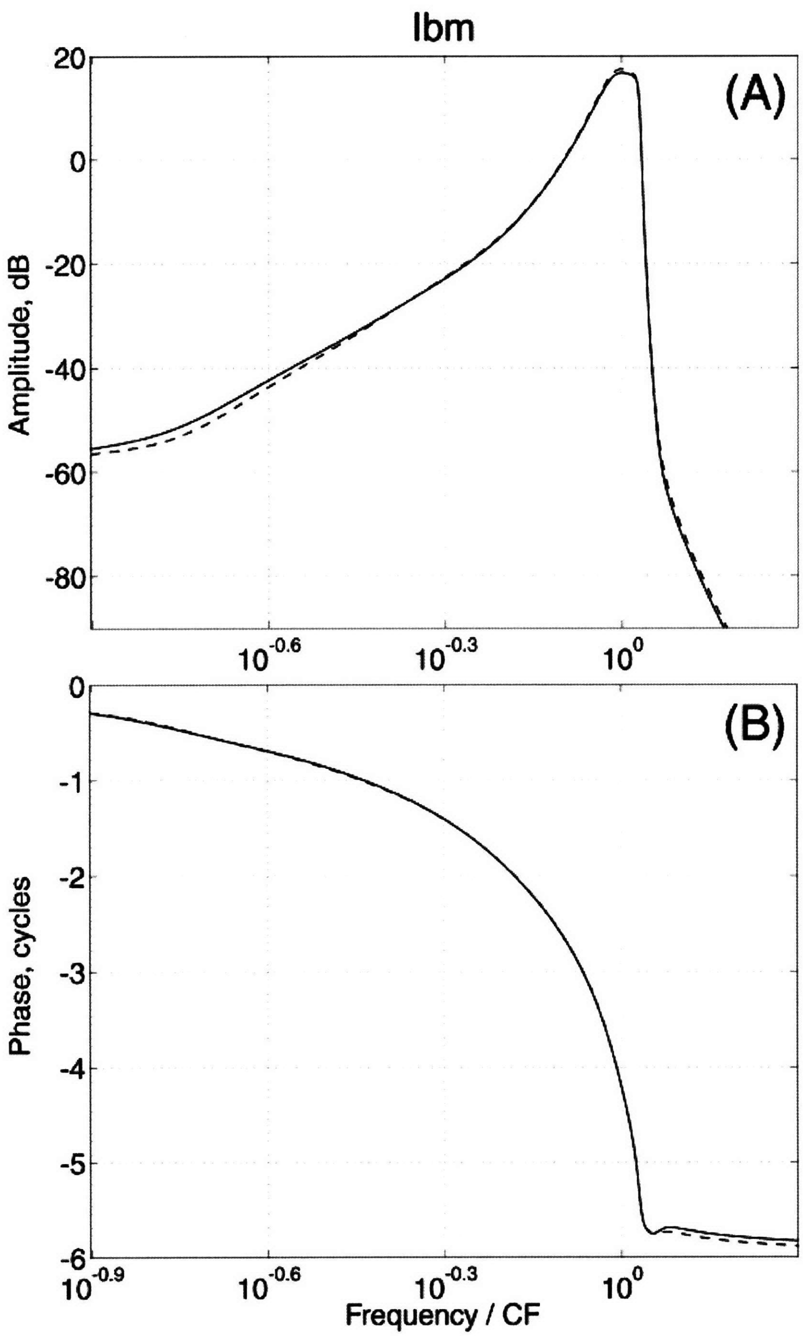


FIG. 8. The local impedance P_l/I_{rl} seen by the first line in the two-mode Sandwich model: (Solid) shows the exact solution, and (Dashed) shows the approximation described in Section II. The top panel shows the real part of the impedance (resistance), and the bottom panel shows the imaginary part (reactance).

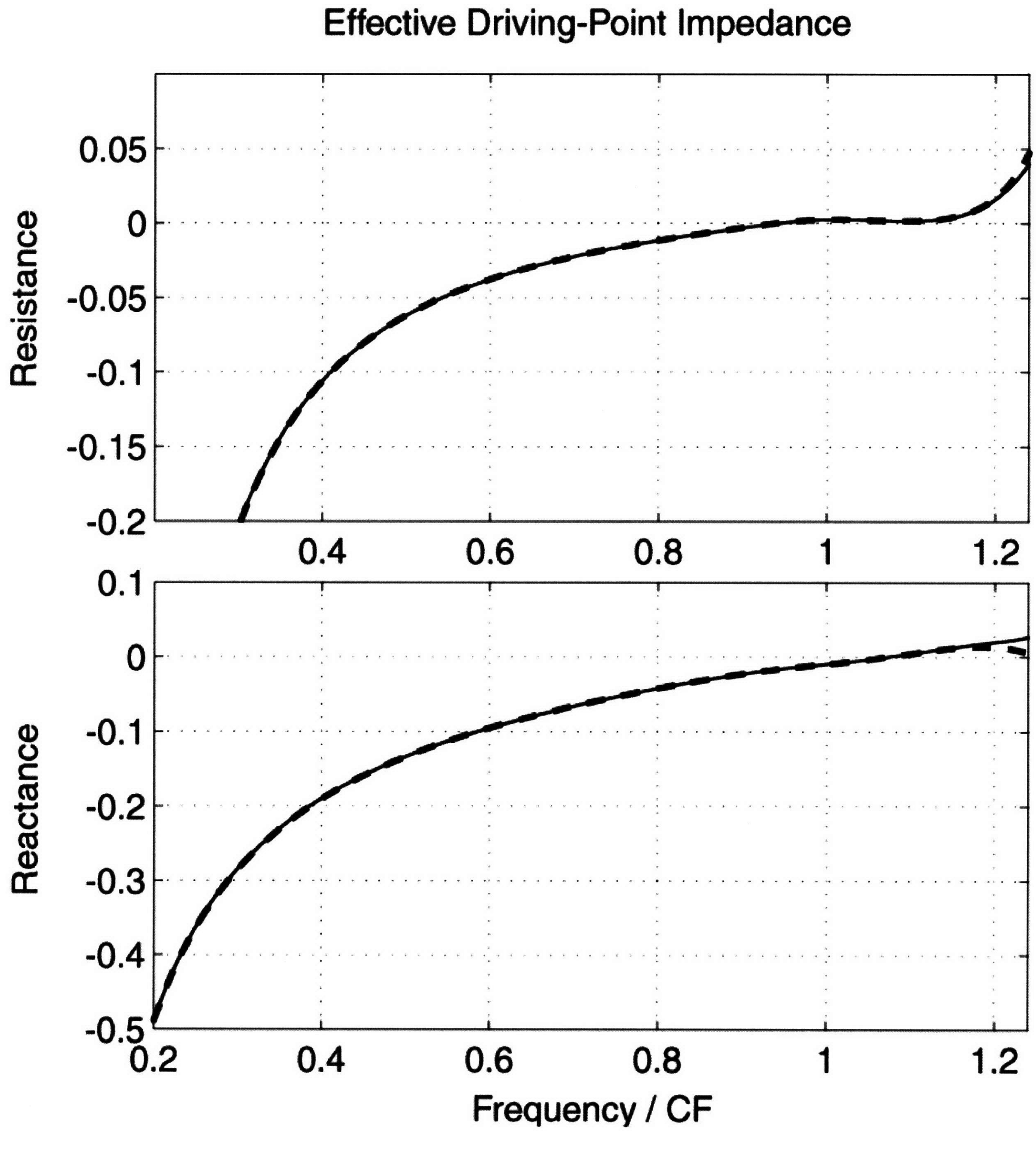


FIG. 9. The two-mode traveling-wave-amplifier cochlear model (Hubbard 1993).

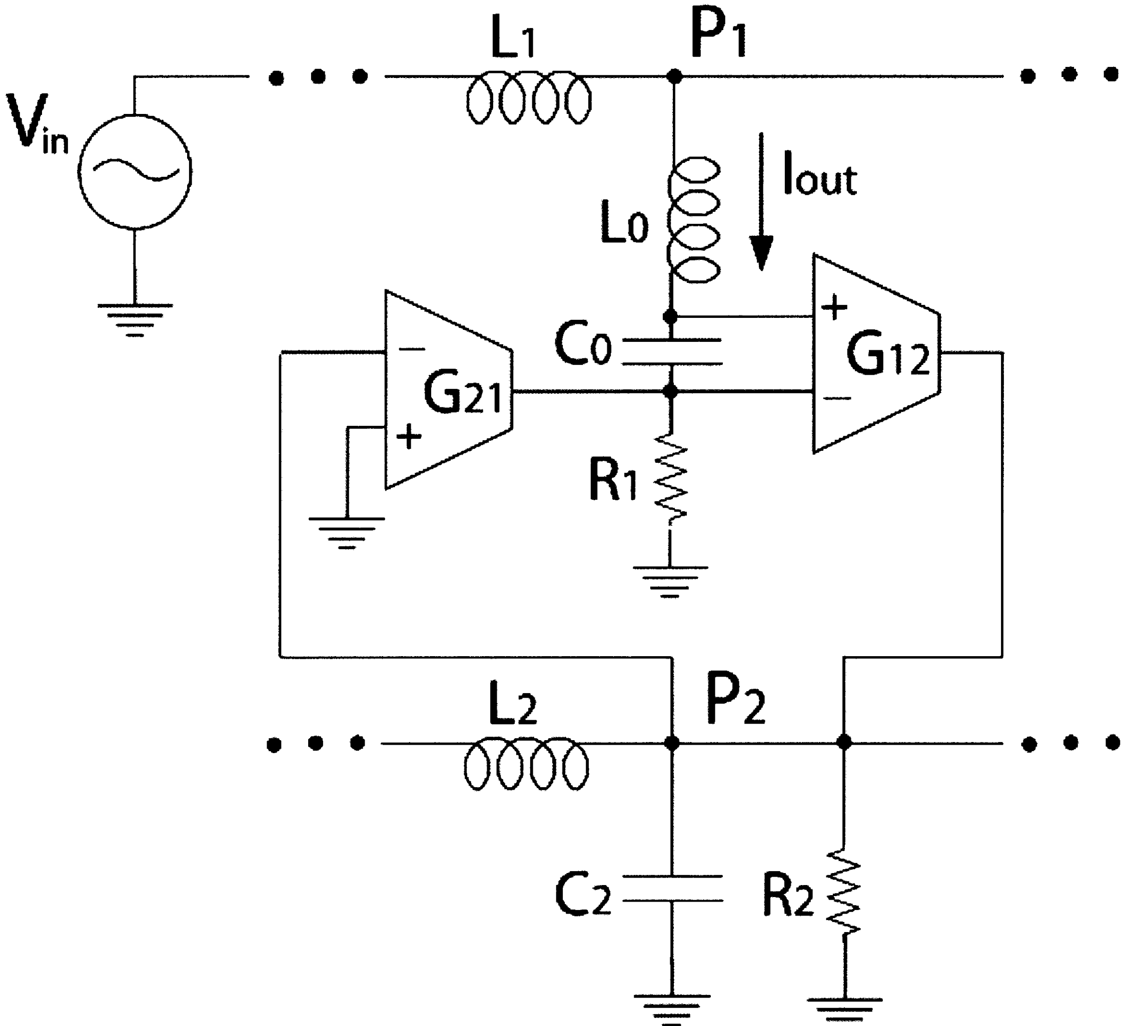


FIG. 10. The pole-zero plot of the rational approximation to $sZ_{eff1,2}(s)$ in the traveling-wave-amplifier cochlear model: (A) The first mode, (B) the second (dominant) mode.

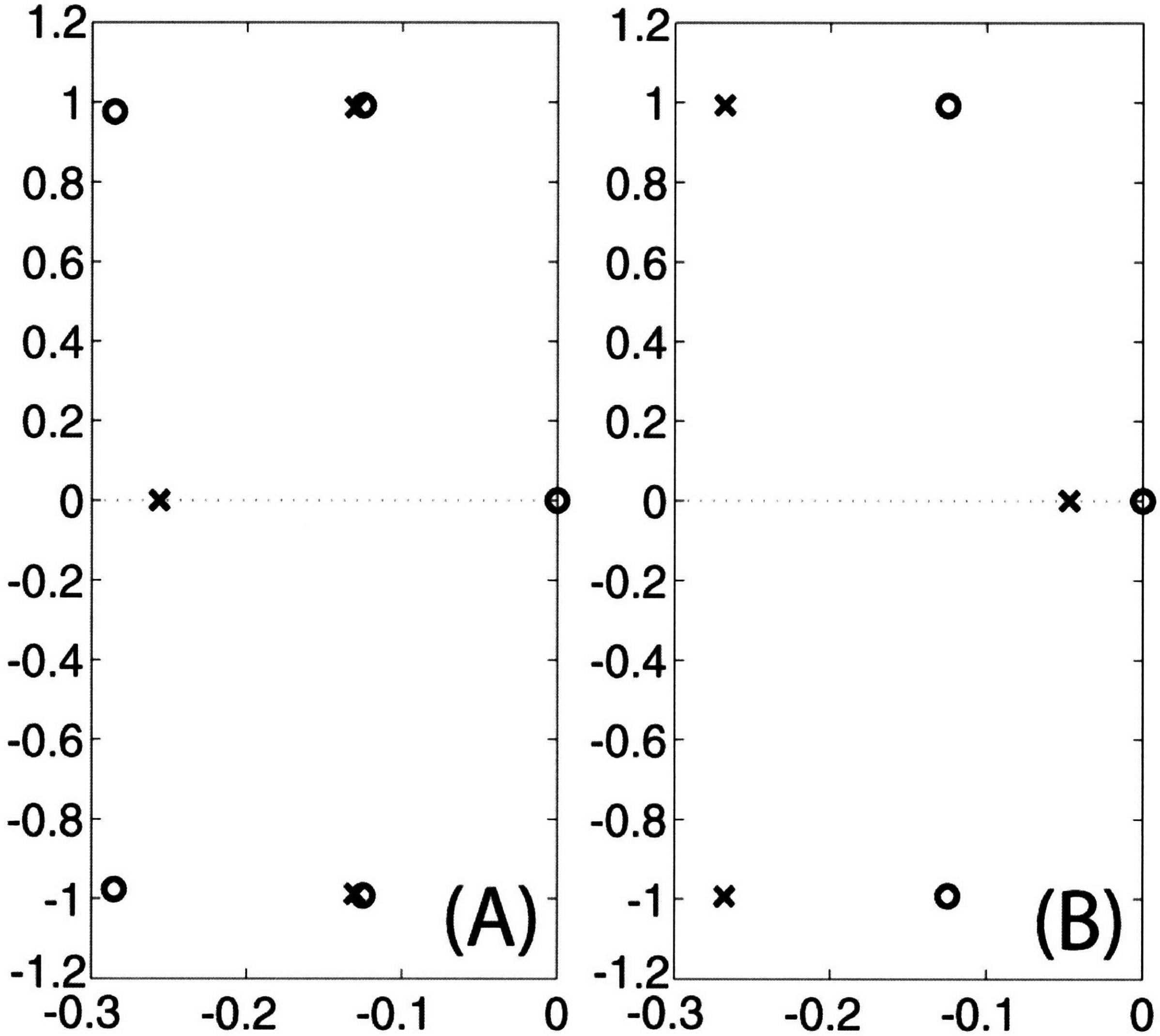


FIG. 11. Modal decomposition of I_{out} amplitude in the traveling-wave-amplifier cochlear model: (Bold solid) shows the exact solution; (Dashed) shows the first mode; (Solid) shows the second mode. Interference between the first and second modes can be seen for *β<0.7*.

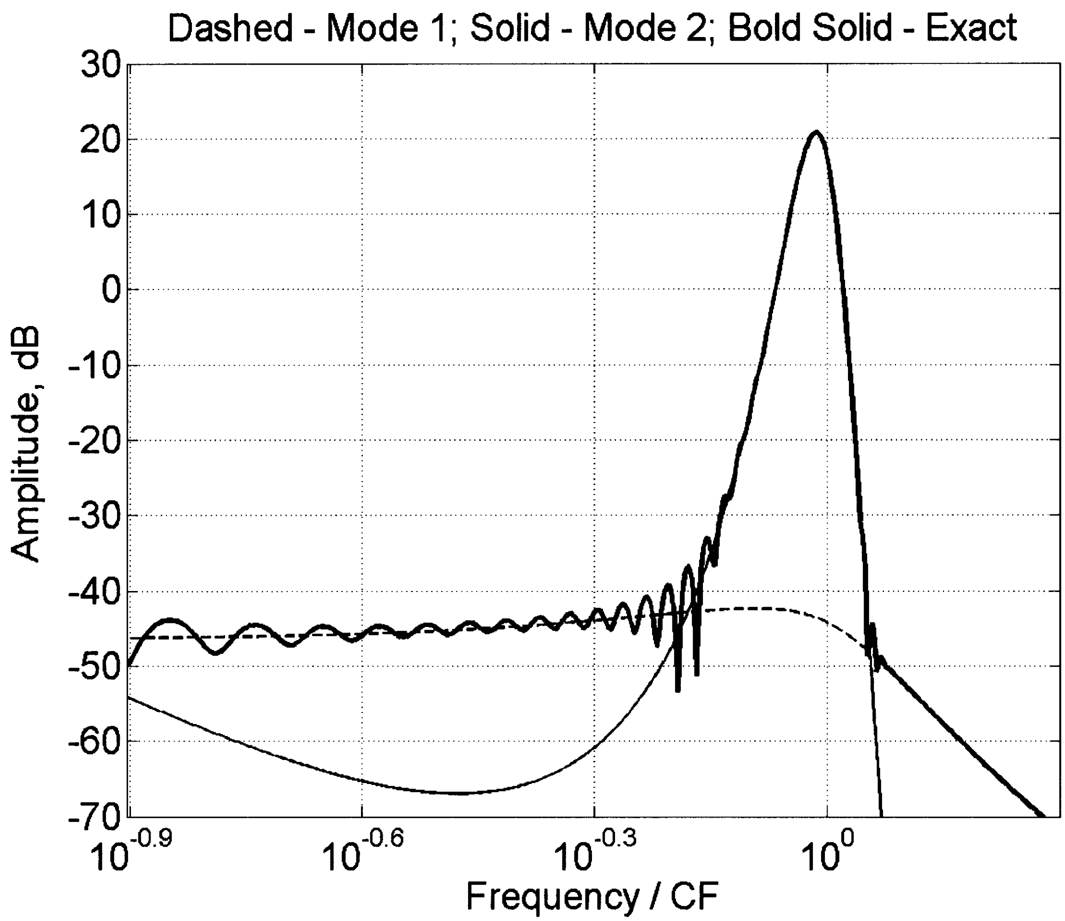


FIG. 12. I_{out} **in the traveling-wave-amplifier cochlear model: (Solid) shows the exact solution; (Dashed) shows the approximation described in Section II; the agreement is excellent. (A) Amplitude (dB), (B) Phase (cycles).**

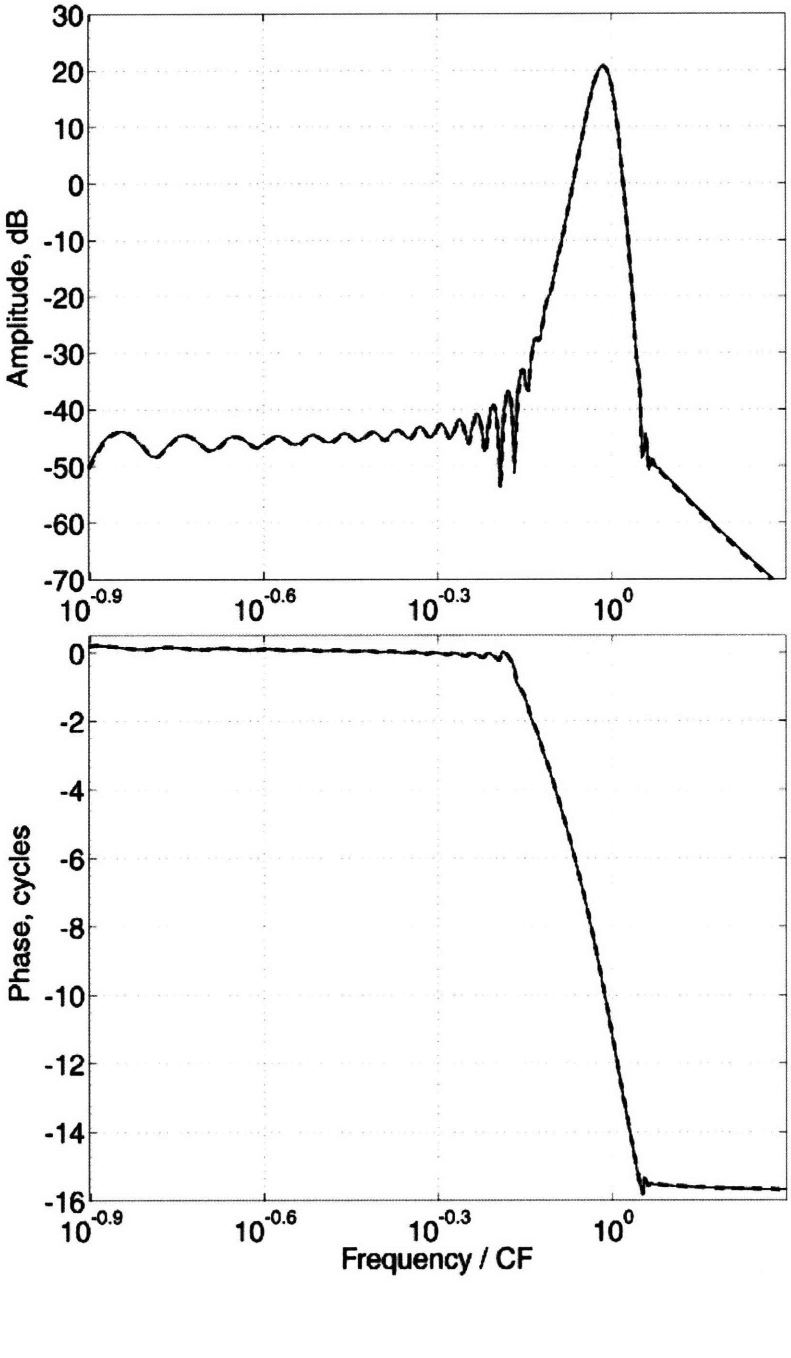


FIG. 13. Modal decomposition of the local impedance P_1/I_{out} seen by the first line in the traveling-wave-amplifier cochlear model: (A) the real part of the impedance (resistance), and (B) the imaginary part (reactance). (Dashed) shows the effective local impedance of the first mode, $Z_{eff1}(s)=1/Y_{eff1}(s)$; (Bold solid) shows the effective local impedance of the second mode, $Z_{eff2}(s)=1/Y_{eff2}(s)$; (Solid) shows the local impedance computed from the exact solution. Note that for $\beta>1.1$, where the first mode dominates, the exact impedance follows the effective impedance of the first mode. In the region $\beta<0.7$ we see the interference between the first and the second modes, with the first mode dominant. For $0.7<\beta<1.1$, where the second mode dominates, the exact impedance follows the effective impedance of the second mode. The real and imaginary parts of this effective impedance are shown with greater resolution in (C).

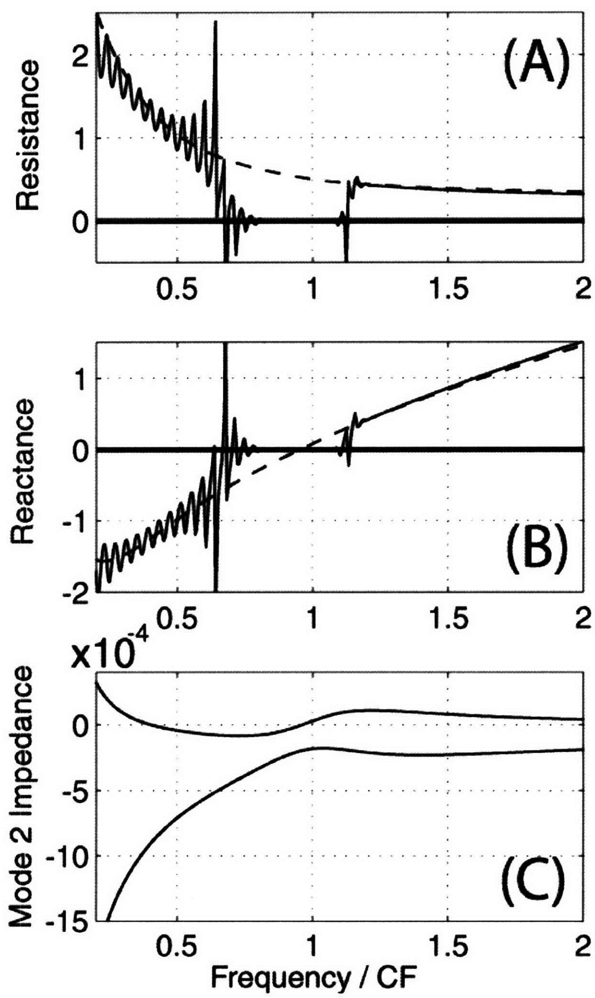
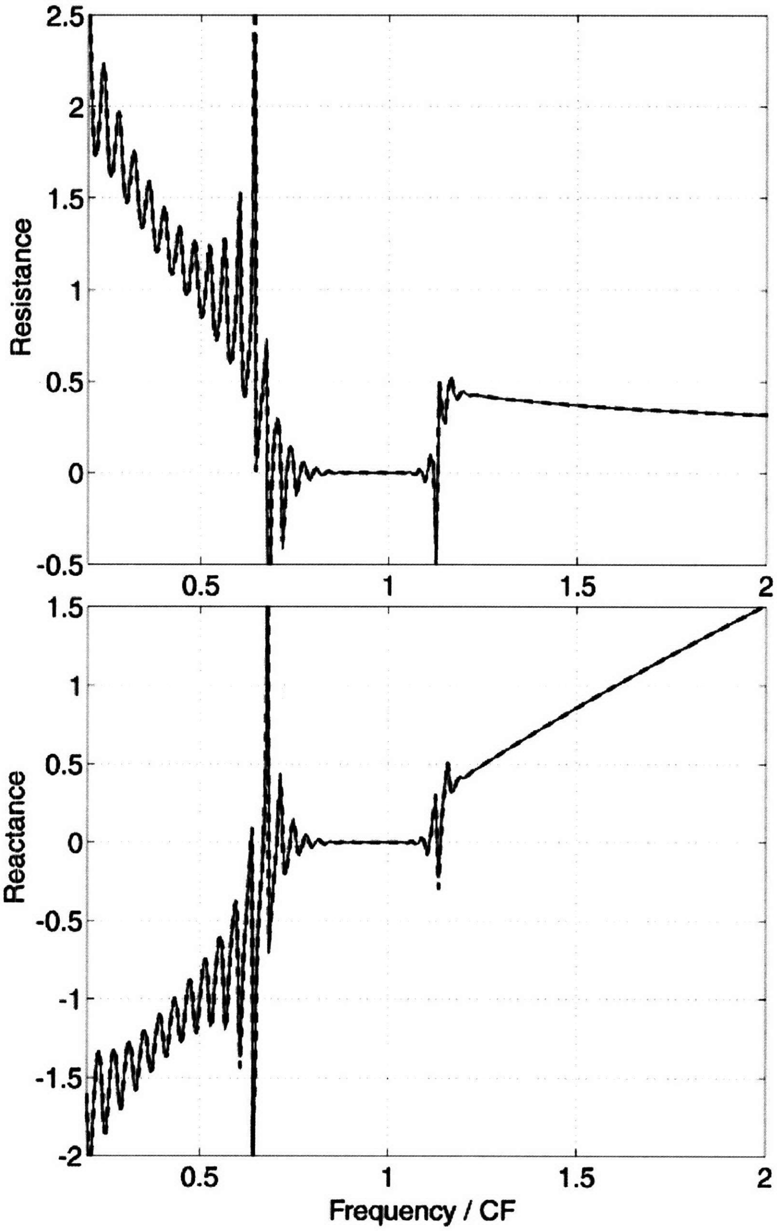


FIG. 14. The local impedance P_1/I_{out} seen by the first line in the traveling-wave-amplifier cochlear model: (Solid) shows the exact solution, and (Dashed) shows the approximation described in Section II. The top panel shows the real part of the impedance (resistance), and the bottom panel shows the imaginary part (reactance). Agreement between our approximate analytical and the exact numerical techniques is excellent.



5. High-Q Low Power Wide Dynamic Range Log-Domain Filter Design

Abstract—A technique that simultaneously reduces the power consumption and increases the SNR and the dynamic range of log-domain filters with high Q is introduced. As an example, a second-order low-pass filter is analyzed, showing the reduction in power consumption and the increase in SNR by a factor of Q. If bias currents in the filter are adjusted as the signal level varies, this technique enables the improvement in maximum SNR by a factor of Q and the increase in maximum non-distorted signal power and the dynamic range by a factor of Q^4. A duality with voltage-mode G_m-C filter design is discussed. Experimental results from a chip in a 0.18-μm 1.1-V CMOS technology are presented for an electronically tunable second-order log-domain filter with adaptive biasing. This filter operates at Q=4, consuming 580-nW at 15-kHz in its quiescent condition. Maximum SNR of 41.3-dB and the dynamic range of 76-dB are achieved. The filter is useful in electronic cochlea, fully implantable bionic ears, hearing aids, and speech-recognition front-ends.

Index Terms—Analog Filters, High-Q, Log-Domain, Adaptive Bias, Figure of Merit, Low Power, Wide Dynamic Range, Bionic Ear, Cochlear Implant

5.1. INTRODUCTION

HIGH resolution frequency discrimination, i.e., high-Q filtering, is required for a variety of applications such as signal processing, speech recognition, hearing aids [1]-[4]. Portable devices are battery powered and required to run off a low voltage, minimize power consumption, and maximize the dynamic range of the system. The challenge in designing biomedical systems is to move to designs that can be fully implanted, and reducing the power consumption is the key. All-analog processing strategies promise power savings of an order of magnitude over even advanced DSP implementations [2]-[4]. However, efficient realization of high-Q analog filters electronically tunable over a wide range of their parameters remains a challenge. The log-domain filtering approach [5]-[7] offers integratable, compact filters that allow wide tuning range and low-voltage operation due to the voltage companding principle [8]. The input signal of the log-domain core is a voltage logarithmically compressed by the diode-connected transistor. The output voltage of the log-domain core is exponentially expanded into current ensuring externally linear operation of the filter. The voltage swings at each node are strongly reduced enabling the low-voltage operation, mitigating parasitic capacitances, and relaxing capacitor linearity requirement. Log-domain can be efficiently combined with dynamic biasing technique [9], where the bias current is kept at the minimum value necessary for the input signal being processed, minimizing noise and power consumption. Log-domain circuits can be realized using CMOS transistors biased in subthreshold. Unlike bipolar transistors, CMOS devices do not suffer from a finite base current. However, threshold voltage mismatches and exiting subthreshold region for higher bias currents limit the performance of CMOS log-domain circuits. This performance degradation is especially pronounced for high-Q filters. Log-domain filters are usually designed using the exponential state-space (ESS) method, in which the desired state-space equations are transformed to the log-domain using an exponential mapping [6], [7]. Related approach is to substitute transconductors in the G_m-C implementation of the filter by nonlinear transconductor blocks. Transfer functions from the input to state-space variables can have amplitudes that are different at DC and at their peak values, usually near the corner frequency. Because of the exponential mapping, no state-space variable can become negative, which means that the maximum amplitude of the signal should not exceed the DC operating point for all state-space variables, otherwise distortion will result. In many filter topologies, both log-domain and linear G_m-C, a single state-space variable becomes a bottleneck, limiting the maximum non-distorted signal and degrading SNR, dynamic range and power consumption. This inefficiency is especially pronounced in high-Q filters, where the disparity between the peak and the DC gain from the input to state-space variables is greatest. The proposed technique adds constant terms to the linear state-space equations, effectively adding DC-biased inputs to shift DC operating points of the state-space variables without altering any transfer functions in the filter. Intuitively, if the DC operating points of all state-space variables are made equal to their signal maximum amplitudes, then all nodes would be biased just as needed and efficiency gains should arise. To judge the efficiency, we utilize the maximum power dissipation of the filter normalized to the 3-dB bandwidth, the order and the maximum SNR, which is a figure of merit (FOM) that can be used to compare filters of different orders and bandwidths [10].

The organization of this paper is as follows. In Section 5.2, we introduce the technique of pre-biasing state-space variables on the example of second-order log-domain low-pass filter design. We analyze noise, maximum non-distorted signal and power consumption effects of pre-biasing. We also quantify the benefits of the proposed technique to log-domain filters with dynamic biasing. In Section 5.3, we compare our log-domain filter to the voltage-mode G_m-C design. In Section 5.4 we present experimental results from a chip. Finally, in section 5.5, we conclude by summarizing the key contributions.

5.2. Theoretical Analysis of Proposed Log-Domain Technique

We would like to implement a second-order low-pass transfer function:

$$TF(s) = \frac{I_{out}}{I_{in}} = \frac{1}{1 + s\tau/Q + s^2\tau^2} \tag{1}$$

Here $s=j\omega$, Q is the quality factor, τ is the time constant, I_{in} and I_{out} are the input and output of the filter, which will be currents in the circuit realization. We would like all parameters to be electronically tunable over wide range of values. State-space realization for this transfer function is not unique, and we pick the following:

$$\begin{aligned} \tau \cdot \dot{x}_1 &= -x_2 + I_{in} \\ \tau \cdot \dot{x}_2 &= x_1 - x_2/Q \\ I_{out} &= x_2 \end{aligned} \tag{2}$$

Here x_1 and x_2 are the state variables. The transfer function from the input to x_1 is:

$$TF_1(s) \triangleq \frac{x_1}{I_{in}} = \frac{s\tau + 1/Q}{1 + s\tau/Q + s^2\tau^2} \tag{3}$$

To implement the log-domain filter in PMOS in weak inversion with each well tied to the respective source terminal, we use the following exponential mappings on the input, output and state variables [7], [11]:

$$x_1 = I_0 \cdot \exp\left(-\frac{\kappa_p \cdot V_1}{U_t}\right); \quad x_2 = I_0 \cdot \exp\left(-\frac{\kappa_p \cdot V_2}{U_t}\right)$$
$$I_{in} = I_0 \cdot \exp\left(-\frac{\kappa_p \cdot V_{in}}{U_t}\right); \quad I_{out} = I_0 \cdot \exp\left(-\frac{\kappa_p \cdot V_{out}}{U_t}\right) \tag{4}$$

Here k_p is the subthreshold exponential parameter of the PMOS transistors, I_0 is some arbitrary constant current, U_t is the thermal voltage kT/q, V_1 and V_2 are the mappings of the state variables, and V_{in} and $V_{out}=V_2$ are the mappings of the input and output that can also be interpreted as input and output voltages of the log-domain core of the filter. Dividing the first equation in (2) by x_1 and the second equation by x_2 and utilizing the mappings (4), we obtain a set of nodal equations that we will realize with PMOS transistors in subthreshold with wells tied to their respective sources, grounded capacitors C, and current sources:

$$-C \cdot \dot{V}_1 = -I_\tau \cdot \exp\left(-\frac{\kappa_p \cdot (V_2 - V_1)}{U_t}\right) + I_\tau \cdot \exp\left(-\frac{\kappa_p \cdot (V_{in} - V_1)}{U_t}\right)$$
$$-C \cdot \dot{V}_2 = I_\tau \cdot \exp\left(-\frac{\kappa_p \cdot (V_1 - V_2)}{U_t}\right) - \frac{I_\tau}{Q} \tag{5}$$

Here $I_\tau = \dfrac{C \cdot U_t}{\kappa_p \cdot \tau}$ is the current that sets the time constant τ, allowing it to be electronically tunable over several orders of magnitude. Figure 1 shows the circuit implementation of nodal equations (5) using blocks as in [7], [11]. The overall gain of this filter is I_2/I_1 and can be tuned in the wide range.

The input current I_{in} is the sum of the DC component I_{DC} and the signal I_{AC}. The amplitude of I_{AC} should not exceed I_{DC} for I_{in} to stay positive. We determine the DC operating point of the circuit from either its state-space equations (2) or the transfer functions (1) and (3) at DC: $x_{1,DC}=I_{DC}/Q;$ $x_{2,DC}=I_{DC}$. The transfer functions (1) and (3) also give us the maximum amplitudes of signals at x_1 and x_2 that occur near the resonant peak of the filter:

$$\begin{aligned} x_{1,\max AC} &\approx I_{AC}\cdot\sqrt{Q^2+1}\approx Q\cdot I_{AC} \\ x_{2,\max AC} &\approx Q\cdot I_{AC} \end{aligned} \tag{6}$$

Here the approximations are acceptable for Qs higher than 2. For the exponential mappings (4) to hold, state variables x_1 and x_2 need to stay positive, i.e., $Q\cdot I_{AC}<I_{DC}/Q$ for the node x_1, and $Q\cdot I_{AC}<I_{DC}$ for the node x_2. Clearly, the state variable x_1 becomes a bottleneck in the system limiting the maximum input non-distorted signal amplitude to $I_{AC}<I_{DC}/Q^2$. The maximum output non-distorted signal amplitude is therefore I_{DC}/Q. For high Q filters this inefficiency is clearly a problem, because high bias current I_{DC} increases power consumption and noise, lowering the maximum SNR. To eliminate bottlenecks in the filter, the amplitudes of the signals at the state variables should just reach their DC operating points as the amplitude of the input signal I_{AC} reaches its bias I_{DC}. The transfer function should remain the same, so according to the equations (6) the maximum amplitudes of the signals at x_1 and x_2 are $Q\cdot I_{DC}$. We modify the state-space equations (2) to shift the DC operating points for the state variables without changing the transfer function by adding constant terms:

$$\begin{aligned} \tau\cdot\dot{x}_1 &= -\left(x_2-Q\cdot I_{DC}\right)+\left(I_{in}-I_{DC}\right) \\ \tau\cdot\dot{x}_2 &= \left(x_1-Q\cdot I_{DC}\right)-\left(x_2-Q\cdot I_{DC}\right)/Q \\ I_{out} &= x_2 \end{aligned} \tag{7}$$

This modification to the state-space representation is equivalent to adding the DC-biased inputs to the filter. Denoting $I_a\equiv I_{DC}\cdot\left(Q-1\right)$ for $Q>1$, and appending the exponential mappings (4) with $I_a=I_0\cdot\exp\left(-\dfrac{\kappa_p\cdot V_a}{U_t}\right)$, we obtain the new set of nodal equations:

$$\begin{aligned} -C\cdot\dot{V}_1 &= -I_\tau\cdot\exp\left(-\frac{\kappa_p\cdot\left(V_2-V_1\right)}{U_t}\right)+I_\tau\cdot\exp\left(-\frac{\kappa_p\cdot\left(V_{in}-V_1\right)}{U_t}\right)+I_\tau\cdot\exp\left(-\frac{\kappa_p\cdot\left(V_a-V_1\right)}{U_t}\right) \\ -C\cdot\dot{V}_2 &= I_\tau\cdot\exp\left(-\frac{\kappa_p\cdot\left(V_1-V_2\right)}{U_t}\right)-\frac{I_\tau}{Q}-I_\tau\cdot\exp\left(-\frac{\kappa_p\cdot\left(V_a-V_2\right)}{U_t}\right) \end{aligned} \tag{8}$$

Figure 2 shows the circuit realization of these nodal equations using blocks as in [7], [11]. In fact, the only modification to the previous implementation is addition of the two PMOS transistors to the log-domain core and the log-compressor transistor, which is highlighted in Figure 2. This filter can also be dynamically biased as in [9]. In this case it needs to be pseudo-differential, and Figure 2 shows only one of the two identical log-domain cores. It is clear that by setting $I_a=0$ the implementation in Figure 2 reduces to that in Figure 1, so we consider only our proposed realization with bias current I_a varying between 0 and $I_{DC}\cdot(Q-1)$.

We shall analyze the power consumption of the log-domain core in Figure 2, ignoring the power associated with I_1, I_2, the input and the output DC currents for now. For convenience, we define a dimensionless variable $Q_a \equiv \frac{I_a}{I_{DC}}+1$, which varies from 1 to Q as we change the bias current I_a from 0 to $I_{DC}\cdot\cdot(Q-1)$. Note that $Q_a=1$ corresponds to the conventional implementation in Figure 1. Assuming that $I_1=I_2$ so that the filter gain is 1, and imagining the PMOS transistor with its gate connected to voltage V_1, the source tied to either sources of M1 and M2 or M3, so that its drain current will be equal to x_1, we apply translinear principle [5] to derive the DC currents in the circuit and its power consumption:

$$
\begin{aligned}
I_{M1}\cdot I_{M6} &= I_{M4}\cdot x_1 \\
I_{M2}\cdot I_{M6} &= I_{M5}\cdot x_1 \\
x_2\cdot I_{M6} &= I_{M7}\cdot x_1 \\
x_1\cdot I_{M12} &= I_{M11}\cdot x_2 \\
I_{M2}\cdot I_{M12} &= I_{M13}\cdot x_2
\end{aligned}
\tag{9}
$$

In DC equilibrium, $I_{M4}+I_{M5}=I_{M7}$, so adding the first two equations in the set (9), substituting the sum into the third equation, and recognizing that $I_{M1,DC}=I_{DC}$ and $I_{M2}=I_a$, we have:

$$x_{2,DC} = I_a + I_{DC} = Q_a\cdot I_{DC} \tag{10}$$

Also at DC, $I_{M11}=I_{M13}+I_\tau/Q$, substituting this equation into the fourth and fifth equation of the set (9), using the equation (10) and recognizing that $I_{M12}=I_\tau$, we obtain after some algebra:

$$x_{1,DC} = I_{DC} \cdot Q_a \cdot \left(1 + \frac{1}{Q} - \frac{1}{Q_a}\right) \tag{11}$$

These results allow us to calculate the current consumption of the log-domain core:

$$I_{DD} = 2 \cdot I_{M7} + I_\tau + 2 \cdot I_{M11} + I_\tau = 2 \cdot I_\tau \cdot \left(1 + \frac{1}{1 + \frac{1}{Q} - \frac{1}{Q_a}} + \left(1 + \frac{1}{Q} - \frac{1}{Q_a}\right)\right) \tag{12}$$

The expression in (12) is minimized by the value of Q_a where $1 + \frac{1}{Q} - \frac{1}{Q_a} = 1$, i.e., $Q_a=Q$. Therefore, the current consumption of the proposed circuit realization is the minimum possible:

$$I_{DD}\left(Q_a = Q\right) = 6 \cdot I_\tau \tag{13}$$

The current consumption of the circuit in Figure 1, where $Q_a=1$, equals:

$$I_{DD}\left(Q_a = 1\right) = 2 \cdot I_\tau \cdot \left(1 + Q + \frac{1}{Q}\right) \tag{14}$$

Therefore, the power reduction from the proposed technique is approximately proportional to Q and becomes significant for Qs higher than 3. If the filter is dynamically biased as in [9], its quiescent power consumption is dominated by that of its log-domain core computed in the equations (12)-(14), because in quiescent condition the currents I_1, I_2 and the input bias current I_{DC} only need to drive parasitic capacitances at the frequencies of interest, while the current I_τ needs to be high enough to drive the capacitors C at those frequencies.

We shall analyze the noise in the log-domain filter in Figure 2. Many authors have contributed to the topic of calculating the noise in externally linear and companding systems [12]-[15]. The noise current spectral density for CMOS device in weak inversion is given by:

$$\overline{i_n^2} = 2q \cdot I_D \cdot \Delta f \tag{15}$$

Here I_D is the drain current of the device. If the transistor is sized such that it is in strong inversion when its drain current is I_D, then the noise current spectral density is

$\overline{i_n^2} = 4kT \cdot \frac{2}{3} g_m \cdot \Delta f \approx \alpha \cdot 2q \cdot I_D \cdot \Delta f$, where g_m is the transconductance of the device in strong inversion. We have defined the dimensionless coefficient α that can be calculated as $\alpha = \frac{4}{3} U_t \frac{g_m}{I_D}$. It equals to $\kappa/0.75$ in weak inversion, which is close to 1 as expected. In the strong inversion, α drops with the device efficiency g_m/I_D. We size the transistors in all current mirrors and current sources such that they are in strong inversion for both matching and the noise reasons. We will arbitrarily assume $\alpha=0.5$ for all current mirrors and current sources in Figure 2.

We first calculate the noise in the log-domain core in Figure 2. Like in voltage-mode G_m-C filters, it is convenient to calculate the equivalent noise contribution of each nonlinear transconductor and then compute the noise of the log-domain core based on the transfer functions to the output.

Transistors M4-M10 and the current source I_τ comprise the first nonlinear transconductor shown in Figure 3. First we note that the movement of voltage at the node V_{m1} does not propagate to the current i_{n1}. This is because $I_{M4}+I_{M5}=I_{M7}$ at the DC operating point, and any deviation of the voltage from the equilibrium at the node V_{m1} changes each of the currents I_{M4}, I_{M5} and I_{M7} by the same factor, thus these changes cancel out at the current output i_{n1}. Therefore, only the drain noise currents of transistors M4, M5, M7 and the strongly inverted transistors of the current mirror M9 and M10 are contributing to the equivalent noise output current i_{n1} of this nonlinear transconductor. Calculating I_{M7} from the third equation of the set (9), recognizing that $I_{M6}=I_\tau$ and using the equations (10), (11) and (15), we compute i_{n1}:

$$\frac{\overline{i_{n1}^2}}{2q \cdot \Delta f} = I_{M7} + \left(I_{M4} + I_{M5}\right) + 2 \cdot \alpha \cdot \left(I_{M4} + I_{M5}\right) = \frac{2 \cdot (1+\alpha) \cdot I_\tau}{1 + \frac{1}{Q} - \frac{1}{Q_a}} \tag{16}$$

Transistors M11-M16 and the current sources I_τ and I_τ/Q comprise the second nonlinear transconductor shown in Figure 4. In this circuit the movement of voltage at the node V_{m2} affects the current i_{n2} because $I_{M11} \neq I_{M13}$. The noise currents that flow directly into the node V_{m2} see the low impedance due to the negative feedback loop M12, M14, and cause negligible movement of the voltage at this node. The noise currents that flow into the drain of M12, however, move the voltage V_{m2} and contribute to the current i_{n2}. These currents are due to the noise of the transistor M12 and the current source I_τ. To compute the transfer coefficient from the noise current flowing into the drain of M12 to the output i_{n2}, we assume that the deviation of the voltage V_{m2} changes each of the currents I_{M11}, I_{M12} and I_{M13} by the factor ε. Then $i_{n2}=I_\tau/Q+\varepsilon \cdot I_{M13}-\varepsilon \cdot I_{M11}=(1-\varepsilon) \cdot I_\tau/Q$. Here we recalled that $I_{M11}=I_{M13}+I_\tau/Q$ at the DC equilibrium. The deviation of the voltage V_{m2} must have been caused by the noise current $I_\tau-\varepsilon \cdot I_{M12}=(1-\varepsilon) \cdot I_\tau$. Comparing the two expressions, we see that the transfer coefficient for the noise current is $1/Q$, so for the noise power it is $1/Q^2$. Calculating I_{M13} from the fifth equation of the set (9) and using the equations (10) and (15), we compute i_{n2}:

$$
\begin{aligned}
\frac{\overline{i_{n2}^2}}{2q \cdot \Delta f} &= \frac{1}{Q^2} \cdot \left(I_{M12} + \alpha \cdot I_\tau\right) + I_{M13} + I_{M11} + 2 \cdot \alpha \cdot I_{M11} + \alpha \cdot \frac{I_\tau}{Q} \\
&= 2 \cdot (1+\alpha) \cdot I_\tau \cdot \left(1 + \frac{1}{Q} - \frac{1}{Q_a}\right) - \frac{I_\tau}{Q} \cdot \left(1 - \alpha - \frac{1+\alpha}{Q}\right) \\
&\approx 2 \cdot (1+\alpha) \cdot I_\tau \cdot \left(1 + \frac{1}{Q} - \frac{1}{Q_a}\right)
\end{aligned}
\tag{17}
$$

We compute the linearized transfer functions from i_{n1} and i_{n2} to the output I_{out}. Calculating the linearized transfer functions for the noise in this nonlinear system is justified because the noise is just a small perturbation around the DC operating point. Rewriting the nodal equations (8), implemented by the proposed log-domain core, and including the currents i_{n1} and i_{n2}, we obtain:

$$
\begin{aligned}
-C \cdot \dot{V}_1 + i_{n1} &= -I_\tau \frac{x_2}{x_1} + I_\tau \frac{I_{in}}{x_1} + I_\tau \frac{I_a}{x_1} \\
-C \cdot \dot{V}_2 + i_{n2} &= I_\tau \frac{x_1}{x_2} - \frac{I_\tau}{Q} - I_\tau \frac{I_a}{x_2}
\end{aligned}
\tag{18}
$$

Multiplying the first equation in (18) by x_1/I_τ and the second equation by x_2/I_τ, and linearizing at the DC operating point, we have:

$$
\begin{aligned}
&\tau \cdot \dot{x}_1 + i_{n1} \cdot \frac{x_{1,DC}}{I_\tau} = -x_2 + I_{DC} + I_a \\
&\tau \cdot \dot{x}_2 + i_{n2} \cdot \frac{x_{2,DC}}{I_\tau} = x_1 - x_2/Q - I_a
\end{aligned} \tag{19}
$$

The linearized transfer functions from i_{n1} and i_{n2} to the output $I_{out}=x_2$ are:

$$
\begin{aligned}
&\frac{x_2}{i_{n1}} = -\frac{x_{1,DC}}{I_\tau} \cdot \frac{1}{1+s\tau/Q+s^2\tau^2}, \quad i_{n2}=0 \\
&\frac{x_2}{i_{n2}} = -\frac{x_{2,DC}}{I_\tau} \cdot \frac{s\tau}{1+s\tau/Q+s^2\tau^2}, \quad i_{n1}=0
\end{aligned} \tag{20}
$$

Using the equations (10), (11), (16), (17) and (20), we calculate the spectral power of the output noise of the filter:

$$
\frac{\overline{i}^2_{n,out}}{2q \cdot \Delta f} = \frac{2 \cdot (1+\alpha) \cdot I_{DC}^2 \cdot Q_a^2}{I_\tau} \cdot \left(1 + \frac{1}{Q} - \frac{1}{Q_a}\right) \cdot \frac{1+|s\tau|^2}{\left|1+s\tau/Q+s^2\tau^2\right|^2} \tag{21}
$$

To get the total output noise over all frequencies, we integrate the spectral power $\overline{i}^2_{n,out}(f)$ given in the equation (21) from 0 to ∞. It can be shown by contour integration that:

$$
\begin{aligned}
&\int_0^\infty \left|\frac{1}{1-x^2+jx/Q}\right|^2 dx = \frac{\pi}{2} Q \\
&\int_0^\infty \left|\frac{jx}{1-x^2+jx/Q}\right|^2 dx = \frac{\pi}{2} Q
\end{aligned} \tag{22}
$$

The total output current noise power over all frequencies due to the log-domain core is then given by:

$$
\overline{i}^2_{tot,out} = \frac{2 \cdot (1+\alpha) \cdot I_{DC}^2 \cdot Q_a^2 \cdot \kappa_p \cdot q}{C \cdot U_t} \cdot \left(Q + 1 - \frac{Q}{Q_a}\right) \tag{23}
$$

Referring this noise current to the output of the log-domain core V_{out} in Figure 2, we obtain the equivalent noise voltage:

$$
\overline{\upsilon}_n^2 = \frac{\overline{i}^2_{tot,out}}{g^2_{m,M3}} = \frac{kT}{C} \cdot \frac{2 \cdot (1+\alpha)}{\kappa_p} \cdot \left(Q + 1 - \frac{Q}{Q_a}\right) \tag{24}
$$

Here $g_{m,M3}$ is the transconductance of the transistor M3 in weak inversion $g_{m,M3}=\kappa_p \cdot I_{M3}/U_t=\kappa_p \cdot x_{2,DC}/U_t=\kappa_p \cdot Q_a \cdot I_{DC}/U_t$, and the last equality uses the equation (10). The noise voltage in the equation (24) is independent from the bias I_{DC}, which allows us to draw the equivalent circuit of the log-domain filter shown in Figure 5 that is valid even for the dynamic biasing. The equation (24) and Figure 5 also imply that one can compute $\overline{\upsilon_n^2}$ treating any log-domain core as a voltage-mode circuit and calculating its small-signal noise transfer functions. Figure 5 suggests the other sources of noise in the system; minimizing their effects leads to additional design considerations. As an example, we study the noise contribution due to the current I_1. The noise current i_{n0} flowing into the drain of the transistor M0 moves the voltage $\upsilon_{01} = \dfrac{i_{n0}}{g_{m,M0}} = \dfrac{i_{n0} \cdot U_t}{I_1 \cdot \kappa_p}$, where we are using small-signal voltage-mode analysis and the subthreshold expression for $g_{m,M0}$. Voltages V_{in} and V_a move in tandem with V_{01}, and these perturbations are filtered by the log-domain core, adding to its output voltage noise $\overline{\upsilon_n^2}$. The small-signal voltage transfer function of the core is $\dfrac{\upsilon_{out}}{\upsilon_{in}} = \dfrac{1/Q_a}{1+s\tau/Q+s^2\tau^2}$, $\upsilon_a = 0$. The factor $1/Q_a$ is due to the ratio of transconductances of the input and the output devices that equals to $I_{DC}/x_{2,DC}$. Using the equation (19) with $i_{n1}=i_{n2}=0$, we determine the transfer function from I_a to the output $I_{out}=x_2$:

$$\frac{x_2}{I_a} = \frac{1-s\tau}{1+s\tau/Q+s^2\tau^2}, \quad I_{in} = 0 \tag{25}$$

The ratio of transconductances of the transistor M2 and the output transistor M3 is $I_a/x_{2,DC}=1-1/Q_a$, so the small-signal voltage transfer function of the core is $\dfrac{\upsilon_{out}}{\upsilon_a} = \left(1-\dfrac{1}{Q_a}\right) \cdot \dfrac{1-s\tau}{1+s\tau/Q+s^2\tau^2}$, $\upsilon_{in} = 0$. Because υ_{in} and υ_a correlate perfectly, we have:

$$\frac{\upsilon_{out}}{i_{n0}} = \frac{U_t}{I_1 \cdot \kappa_p} \cdot \frac{1-s\tau \cdot (1-1/Q_a)}{1+s\tau/Q+s^2\tau^2} \tag{26}$$

From the equation (15) we see that $\frac{\overline{i_{n0}^2}}{2q\cdot\Delta f}=I_1\cdot(1+\alpha)$. Then taking the square of the absolute value of the equation (26), integrating over all frequencies and applying the equations (22), we calculate the additional voltage noise contribution due to the current I_1 in Figure 5:

$$\overline{\upsilon_{n1}^2}=\frac{U_t^2}{\kappa_p^2}\cdot\frac{q}{\tau\cdot I_1}\cdot\frac{1+\alpha}{2}\cdot Q\cdot\left[1+\left(1-\frac{1}{Q_a}\right)^2\right] \tag{27}$$

We need to ensure that this additional noise does not significantly degrade filter's performance, i.e., that $\overline{\upsilon_{n1}^2}<\overline{\upsilon_n^2}$. This condition leads to the following design consideration for the current I_1:

$$I_1>\frac{I_\tau}{4}\cdot\frac{1+(1-1/Q_a)^2}{1+1/Q-1/Q_a} \tag{28}$$

For the circuit in Figure 1, where $Q_a=1$, the condition (28) yields:

$$I_1>Q\cdot I_\tau/4 \tag{29}$$

For the proposed technique, where $Q_a=Q$, the condition (28) implies:

$$I_1>I_\tau/2 \tag{30}$$

Therefore, the power consumption of the control circuits in the proposed circuit of Figure 2 can be reduced by a factor of $Q/2$. In the equations (12)-(14) we obtained the similar power savings from the log-domain core.

Finally, the input transistors M1 and M2, and the output transistor M3 contribute the noise of their own. Using the equation (15) and the transfer functions (1) and (25), we compute the uncorrelated noise contributions of the input devices M1 and M2. Integrating over all frequencies, we obtain:

$$\overline{i_{tot,cm}^2}=\frac{(1+\alpha)\cdot q\cdot Q}{\tau}\cdot\left(Q_a-\tfrac{1}{2}\right)\cdot I_{DC} \tag{31}$$

Factor α accounts for the fact that current sources supply the input currents I_{in} and I_a. Note that the noise power component (31) is proportional to the bias current I_{DC}, not to I_{DC}^2. This dependency is analogous to the noise of the current mirror. For large bias currents I_{DC} the contribution (31) is negligible. However, as we reduce I_{DC} in the dynamically biased filters, the noise contribution (31) decreases slower than that in the equation (23) and might become dominant. Further reduction in I_{DC} in this case does not lead to the improvements in filter's noise performance expected from the equation (23).

If the conditions (28)-(30) for the current I_l are met and the input bias current I_{DC} is large enough so that the noise term (31) can be neglected, then the equation (23) determines the noise performance of the filter. For the circuit in Figure 1, where $Q_a=1$, the equation (23) gives:

$$\overline{i}^2_{tot,out,1} = \frac{2\cdot(1+\alpha)\cdot\kappa_p\cdot q}{C\cdot U_t}\cdot I^2_{DC} \tag{32}$$

For the proposed technique, where $Q_a=Q$, the equation (23) yields:

$$\overline{i}^2_{tot,out,2} = \frac{2\cdot(1+\alpha)\cdot\kappa_p\cdot q}{C\cdot U_t}\cdot I^2_{DC}\cdot Q^3 \tag{33}$$

The discussion after the equation (6) established that for the conventional circuit in Figure 1, the maximum non-distorted output signal amplitude was I_{DC}/Q, while for the proposed circuit in Figure 2 the maximum non-distorted signal amplitude was $I_{DC}\cdot Q$. Therefore, the maximum SNR, or the dynamic range, of the filter in Figure 1, is:

$$SNR_{\max,1} \equiv \frac{I^2_{\max AC,1}/2}{\overline{i}^2_{tot,out,1}} = \frac{C\cdot U_t}{4\cdot(1+\alpha)\cdot\kappa_p\cdot q}\cdot\frac{1}{Q^2} \tag{34}$$

The maximum SNR, or the dynamic range, of the proposed filter in Figure 2, is:

$$SNR_{\max,2} \equiv \frac{I^2_{\max AC,2}/2}{\overline{i}^2_{tot,out,2}} = \frac{C\cdot U_t}{4\cdot(1+\alpha)\cdot\kappa_p\cdot q}\cdot\frac{1}{Q} \tag{35}$$

So, the dynamic range of the proposed filter is improved by a factor Q over the conventional log-domain filter topology. This improvement is in addition to the power savings (12)-(14) proportional to Q in the log-domain core, and the similar power reduction (28)-(30) in the control circuitry. To quantify this efficiency, we use the figure of merit (FOM), which is the power consumption of the filter normalized to its 3-dB bandwidth, the order and the maximum SNR [10]. The equations (14) and (34) yield the FOM of the filter in Figure 1:

$$FOM_1 \equiv \frac{V_{DD} \cdot I_{DD,1} \cdot \tau/Q}{2 \cdot SNR_{\max,1}} = 4 \cdot (1+\alpha) \cdot q \cdot V_{DD} \cdot (Q^2 + Q + 1) \tag{36}$$

We calculate the FOM of the proposed filter in Figure 2 using the equations (13) and (35):

$$FOM_2 \equiv \frac{V_{DD} \cdot I_{DD,2} \cdot \tau/Q}{2 \cdot SNR_{\max,2}} = 12 \cdot (1+\alpha) \cdot q \cdot V_{DD} \tag{37}$$

We see that the efficiency, expressed as the FOM, is improved by a factor $(Q^2+Q+1)/3$ by the proposed technique. Even for $Q=4$ this improvement amounts to a factor 7, and grows quadratically for higher Qs. It is worth mentioning, however, that we have excluded the input bias I_{DC} and the output $Q_a \cdot I_{DC}$ currents from the power consumption calculation in the equations (36) and (37), which will diminish the benefit of the proposed technique somewhat if the bias currents are not dynamically adjusted. In this case we recommend reducing the overall gain of the filter by either sizing down the transistors M2 and M3 or decoupling the reference voltages V_{ref} and adjusting each one separately as recommended in [11]. In the dynamically biased filters the quiescent power consumption of the input and output devices is negligible because the minimum I_{DC} only needs to drive device parasitic capacitances and can therefore be much smaller than I_τ that drives capacitors C.

As we discuss next, the efficiency improvement enabled by the proposed technique for log-domain filters can also be observed in voltage-mode filters. Selecting the topologies carefully, we can achieve the similar efficiencies for high Qs in the G_m-C designs.

5.3. Comparison to Voltage-Mode G_m-C Design

A conventional voltage-mode second-order low-pass filter topology was analyzed in [2] and is shown in Figure 6 for convenience. It was derived in [2] that this filter implements the transfer function (1) if $G_1=(C/\tau)\cdot Q$ and $G_2=(C/\tau)/Q$. If the transconductors with the input voltage swing V_L

are used in Figure 6, their bias currents are $I_1=G_1 \cdot V_L$ and $I_2=G_2 \cdot V_L$ as described in [16]. The current consumption of the transconductor circuit in [16] and [2] is approximately twice its bias current; therefore, the current consumption of the filter in Figure 6 is:

$$I_{DD,3} = 2 \cdot \frac{C \cdot V_L}{\tau} \cdot \left(Q + \frac{1}{Q} \right) \tag{38}$$

We determine the maximum undistorted output signal by assuming that the distortion of the transconductor is low as long as the voltage between its positive and negative input terminals does not exceed V_L. The transfer functions from the input voltage V_{in} to the voltages V_{diff1} and V_{diff2} between the first and the second transconductors' input terminals in Figure 6 are:

$$\begin{aligned} \frac{V_{diff1}}{V_{in}} &= \frac{s\tau/Q + s^2\tau^2}{1 + s\tau/Q + s^2\tau^2} \\ \frac{V_{diff2}}{V_{in}} &= \frac{s\tau \cdot Q}{1 + s\tau/Q + s^2\tau^2} \end{aligned} \tag{39}$$

We see that the maximum gain from the input to the first transconductor is approximately $\sqrt{Q^2+1} \approx Q$, and the maximum gain from the input to the second transconductor is about Q^2. Therefore, saturating the second transconductor is the bottleneck in the filter with high Q. The maximum undistorted input voltage is thus equal to V_l/Q^2, which gives the maximum undistorted output signal amplitude as $V_{max,out,3}=V_l/Q$. The total output voltage noise power over all frequencies for this filter topology was computed in [2] and is repeated here for convenience:

$$\overline{v}^2_{tot,out,3} = \frac{N \cdot q \cdot V_L}{2 \cdot C} \tag{40}$$

Here N is the effective number of shot-noise sources in the transconductor. Details of how to compute the effective number of noise sources in transconductance circuits are provided in [16]. The maximum SNR, or dynamic range, of this filter and its FOM are then given by:

$$SNR_{\max,3} \equiv \frac{V^2_{\max,out,3}/2}{\bar{\upsilon}^2_{tot,out,3}} = \frac{C \cdot V_L}{N \cdot q} \cdot \frac{1}{Q^2}$$
$$FOM_3 \equiv \frac{V_{DD} \cdot I_{DD,3} \cdot \tau/Q}{2 \cdot SNR_{\max,3}} = N \cdot q \cdot V_{DD} \cdot \left(Q^2+1\right) \tag{41}$$

The equations (38), (40) and (41) for the voltage-mode filter in Figure 6 have very similar structure to the equations (14), (32), (34) and (36) for the conventional log-domain filter in Figure 1, if we substitute N by $4{\cdot}(1+\alpha)$, and V_L by U_t/κ_p. In fact, the input voltage swing of a simple 5-transistor differential-pair transconductor is $V_L=U_t/\kappa_p$, and typical values of N are 4.8-5.3, which is very close to $4{\cdot}(1+\alpha)$ for typical values of α as was illustrated in [16]. Both topologies are similarly inefficient for high-Q filter realization, and we would like to design a voltage-mode circuit that improves the efficiency for high Qs mirroring our proposed technique for the log-domain filters. We generally start with the state-space representation with its coefficients either independent or inversely proportional to Q. In voltage-mode G_m-C realizations the state-space variables are the voltages on the grounded capacitors and the coefficients are proportional to the transconductances and hence the power consumption. The state-space representations (2) or (7) satisfy the above requirement, and their G_m-C realizations are the same. Figure 7 shows the half of the fully differential voltage-mode circuit implementing the state-space equations (2) or (7) and the transfer function (1) if $G=C/\tau$. If the transconductors with the input voltage swing V_L are used, the current consumption of this filter is:

$$I_{DD,4} = 2 \cdot \frac{C \cdot V_L}{\tau} \cdot \left(3 + \frac{1}{Q}\right) \tag{42}$$

The gains from the input V_{in} to the transconductors are determined by the transfer functions $V_{out}/V_{in}(s)$ and $V_1/V_{in}(s)$, given by the equations (1) and (3), respectively. The maximum gains of these transfer functions were calculated in the equations (6); they approximately equal Q. Therefore, the maximum undistorted input voltage is equal to V_l/Q, which gives the maximum undistorted output signal amplitude as $V_{max,out,4}=V_L$.

We shall calculate the output noise of the filter in Figure 7. Similar to the equation (15), each transconductor produces the noise current with spectral density $\overline{i_n^2}/\Delta f = N \cdot q \cdot I_{bias}$ on its output. The calculation of the transfer functions from those noise sources to the output voltage V_{out} is very similar to that of the equations (19), and the result is identical to the equations (20) to the scaling factor. The spectral power of the output noise of the filter is thus given by:

$$\frac{\overline{\upsilon_{n,out}^2}}{\Delta f} = \frac{\tau \cdot N \cdot q \cdot V_L}{C} \cdot \frac{2 + |s\tau|^2 \cdot (1 + 1/Q)}{\left|1 + s\tau/Q + s^2\tau^2\right|^2} \tag{43}$$

Using the integrals (22), we compute the total output voltage noise power over all frequencies:

$$\overline{\upsilon_{tot,out,4}^2} = \frac{N \cdot q \cdot V_L \cdot (3 \cdot Q + 1)}{4 \cdot C} \tag{44}$$

The maximum SNR, or dynamic range, of this filter and its FOM are then given by:

$$\begin{aligned} SNR_{\max,4} &\equiv \frac{V_{\max,out,4}^2/2}{\overline{\upsilon_{tot,out,4}^2}} = \frac{C \cdot V_L}{N \cdot q} \cdot \frac{2}{3 \cdot Q + 1} \\ FOM_4 &\equiv \frac{V_{DD} \cdot I_{DD,4} \cdot \tau/Q}{2 \cdot SNR_{\max,4}} = N \cdot q \cdot V_{DD} \cdot \frac{1}{2} \cdot \left(3 + \frac{1}{Q}\right)^2 \end{aligned} \tag{45}$$

The equations (42) and (45) for the G_m-C filter in Figure 7 have very similar form to the equations (13), (35) and (37) for the proposed log-domain filter in Figure 2, if we again substitute N by $4 \cdot (1+\alpha)$, and $C \cdot V_L$ by $C \cdot U_t/\kappa_p$. The advantage of voltage-mode realizations with wide-linear-range transconductors [16] is reducing the size of the capacitance C by a factor of $V_L/(U_t/\kappa_p)$ for a given specification of SNR_{max} and Q, which can be very important in applications like electronic cochlea, fully implantable bionic ears, hearing aids, and speech-recognition front-ends [1]-[4]. The advantage of log-domain filters is revealed when we compare the equation (44) for the total output noise of the G_m-C circuit to the equation (33) for the total output noise of the log-domain design. The former is constant and independent of the signal level, whereas the latter is proportional to the square of the input bias current and thus can be made proportional to the signal power if the bias is dynamically adjusted as in [9]. Therefore, the noise is reduced for the faint signals, keeping the SNR constant at its maximum value over very wide range of input signals thereby increasing the circuit's dynamic range, and reducing the quiescent power consumption.

5.4. Experimental Results

A chip with this filter was fabricated on UMC's 0.18-μm CMOS process. Power supply is 1.1-V. Figures show experimental data taken from the chip.

5.5. Conclusions

The.

5.6. References

[1] W. A. Serdijn, A. C. van der Woerd, J. Davidse, H. M. van Roemund, "A Low-Voltage Low-Power Fully-Integratable Front-End for Hearing Instruments," *IEEE Trans. Circuits Syst. I*, vol. 42 (11), pp. 920-932, Nov. 1995.

[2] 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.

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

[4] R. Sarpeshkar, C. Salthouse, Ji-Jon Sit, M.W. Baker, S.M. Zhak, T.K.-T. Lu, L. Turicchia and S. Balster, "An ultra-low-power programmable analog bionic ear processor," *IEEE Transactions on Biomedical Engineering*, vol. 52 (4), pp. 711-727, Apr. 2005.

[5] B. Gilbert, "Translinear Circuits: A Proposed Classification," *Electronics Letters,* vol. 11, no. 1, pp. 14–16, 1975.

[6] D. R. Frey, "Log-domain filtering: an approach to current-mode filtering," *IEE Proceedings, Part G*, vol. 140, pp. 406-416, 1993.

[7] D. R. Frey, "Log Domain Filtering for RF Applications," *IEEE Journal of Solid-State Circuits*, vol. 31 (10), pp. 1468-1475, Oct. 1996.

[8] Y. Tsividis, V. Gopinathan, and L. Toth, "Companding in signal processing," *Electron. Lett.*, vol. 26, pp. 1331-1332, Aug. 1990.

[9] N. Krishnapura and Y. Tsividis, "Noise and Power Reduction in Filters Through the Use of Adjustable Biasing," *IEEE Journal of Solid State Circuits,* vol. 36 (12), pp. 1912-1920, Dec. 2001.

[10] E. Vittoz *et al.*, "Low power low-voltage limitations and prospects in analog design," in *Analog Circuit Design, Low-Power, Low-Voltage, Integrated Filters and Smart-Power*, R. J. V. D. Plassche *et al.*, Eds. Boston, MA: Kluwer 1995.

[11] D. Python and C. Enz, "A Micropower Class-AB CMOS Log-Domain Filter for DECT Applications," *IEEE Journal of Solid-State Circuits*, vol. 36 (7), pp. 1067-1075, July 2001.

[12] W. A. Serdijn, M. Kouwenhoven, J. Mulder and A. van Roermund, "Design of High Dynamic Range Fully Integratable Translinear Filters," *Analog Integrated Circuits and Signal Processing*, vol. 19, pp. 223-239, 1999.

[13] J. Mulder, M. Kouwenhoven, W. A. Serdijn, A. van der Woerd, and A. van Roermund, "Nonlinear Analysis of Noise in Static and Dynamic Translinear Circuits," *IEEE Tran. Circuits and Syst. II: Analog Digital Sign. Proc.*, vol. 46 (3), pp. 266-278, Mar. 1999.

[14] L. Toth, G. Efthivoulidis, and Y. Tsividis, "Noise Analysis of Externally Linear Systems," *IEEE Tran. Circuits and Syst. II: Analog Digital Sign. Proc.*, vol. 47 (12), pp. 1365-1377, Dec. 2000.

[15] A. Ng, and J. Sewell, "Direct Noise Analysis of Log-Domain Filters," *IEEE Tran. Circuits and Syst. II: Analog Digital Sign. Proc.*, vol. 49 (2), pp. 101-109, Feb. 2002.

[16] 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, pp. 123-151, May/June 1997.

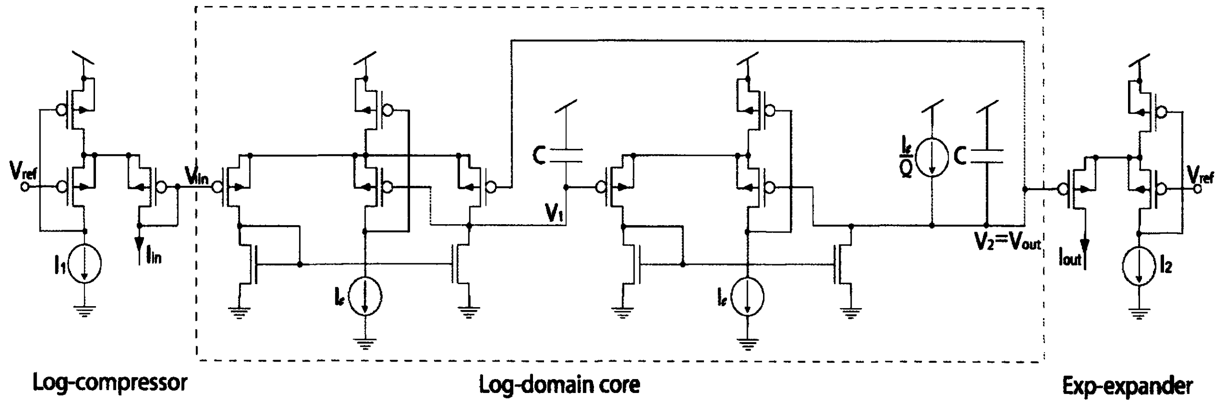


Figure 1: Conventional second-order low-pass log-domain filter topology.

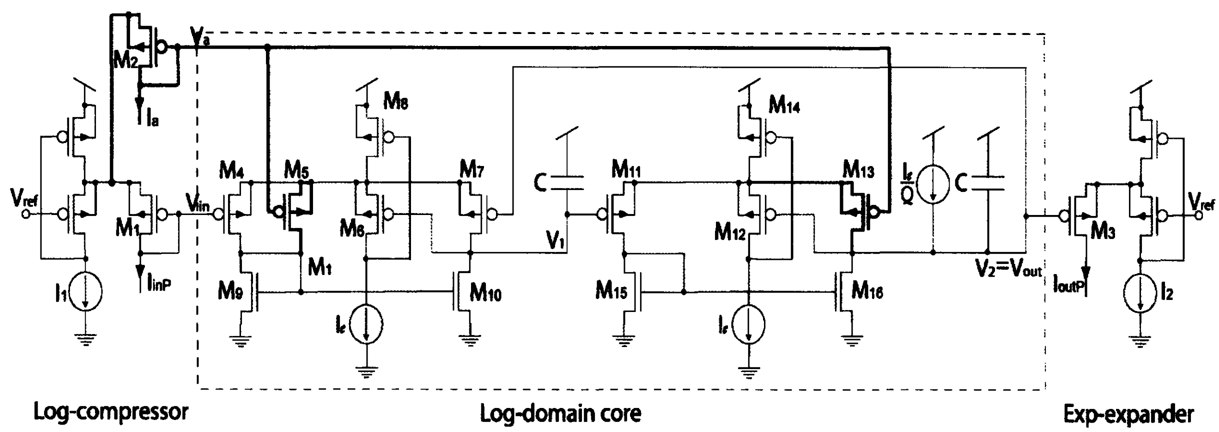


Figure 2: Proposed filter (one-half only is shown). The modifications are highlighted.

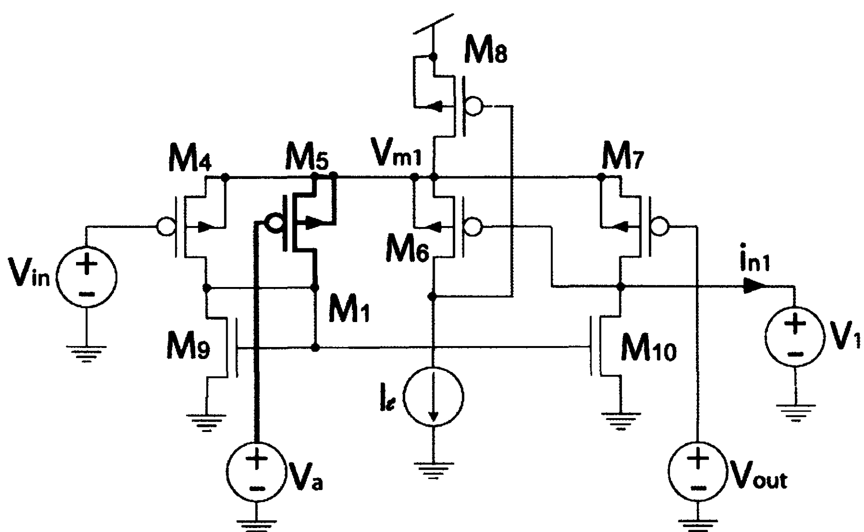


Figure 3: The first nonlinear transconductor from the log-domain core of the filter.

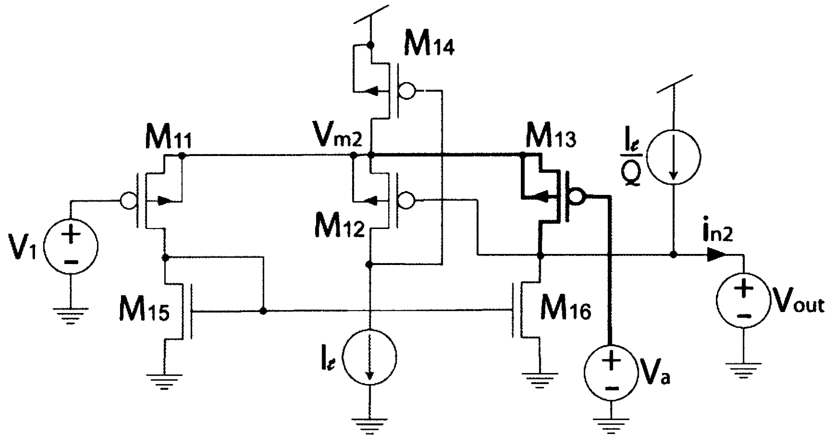


Figure 4: The second nonlinear transconductor from the log-domain core of the filter.

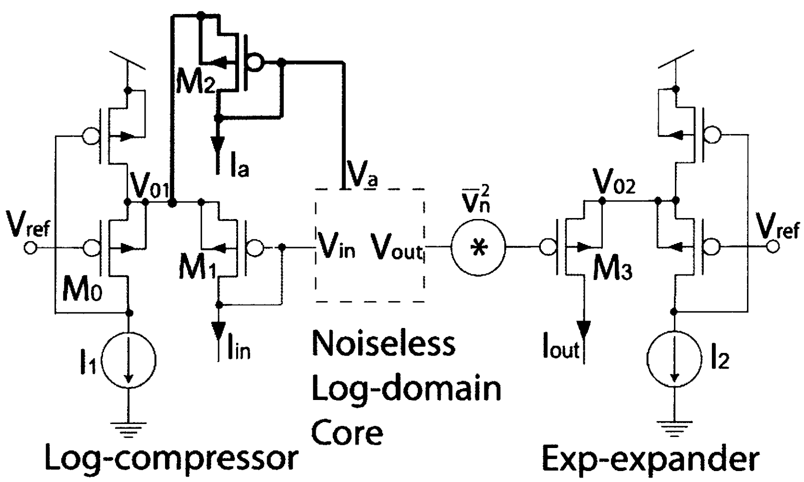


Figure 5: The equivalent circuit of the log-domain filter.

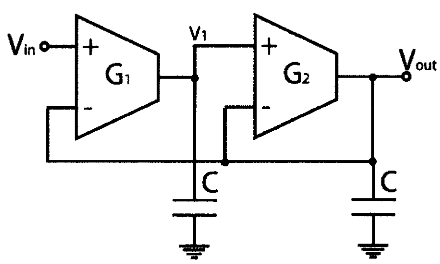


Figure 6: Conventional second-order voltage-mode low-pass filter [2].

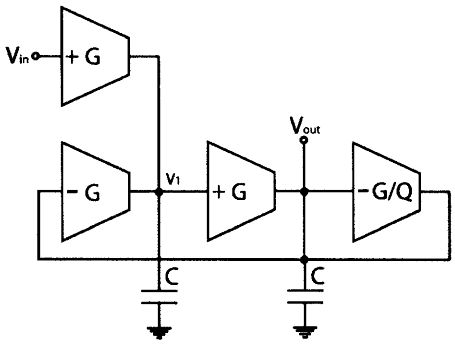


Figure 7: Improved filter modified to become differential (one-half only is shown).

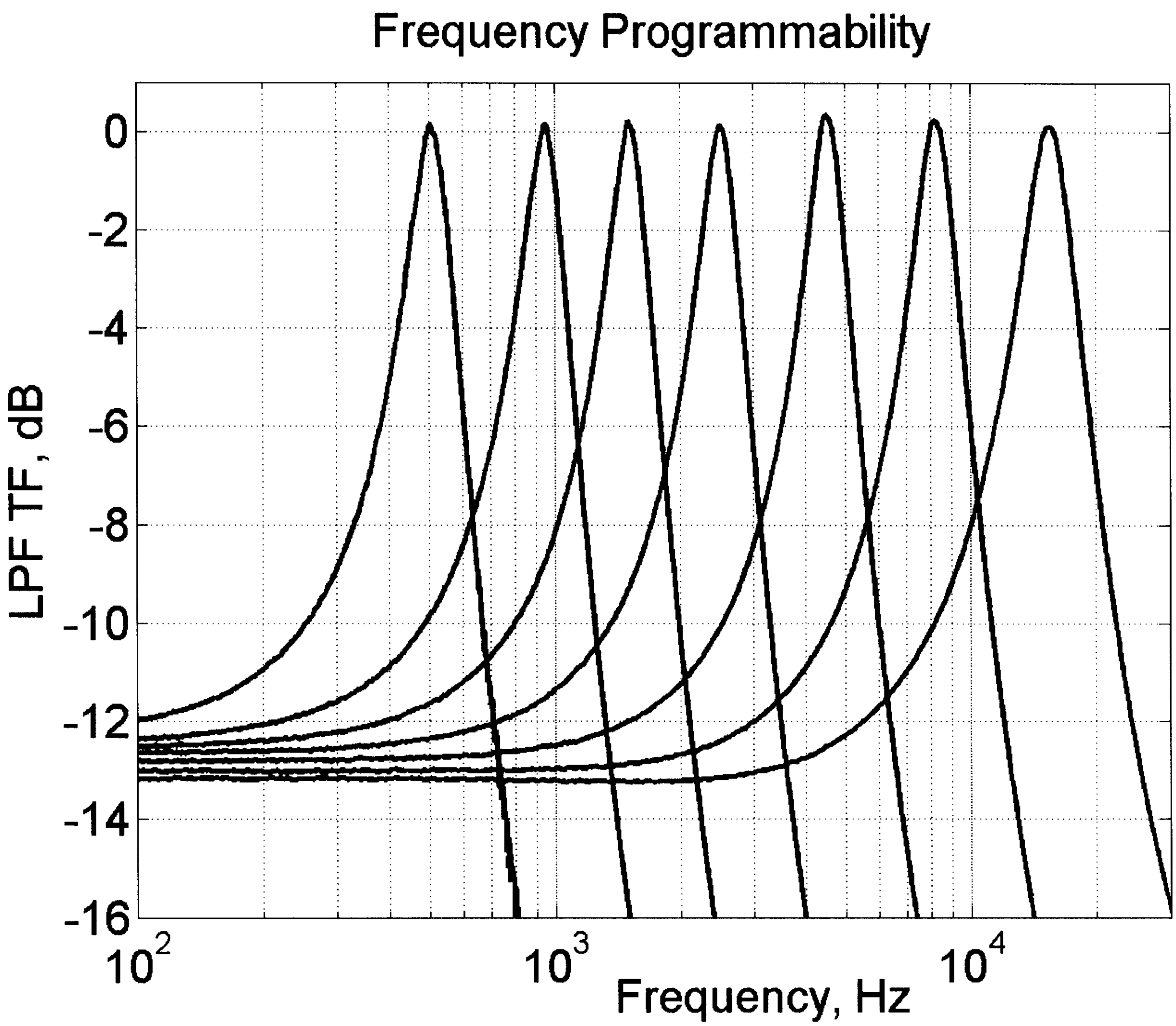


Figure 8: Measured transfer functions for various frequency settings of the filter.

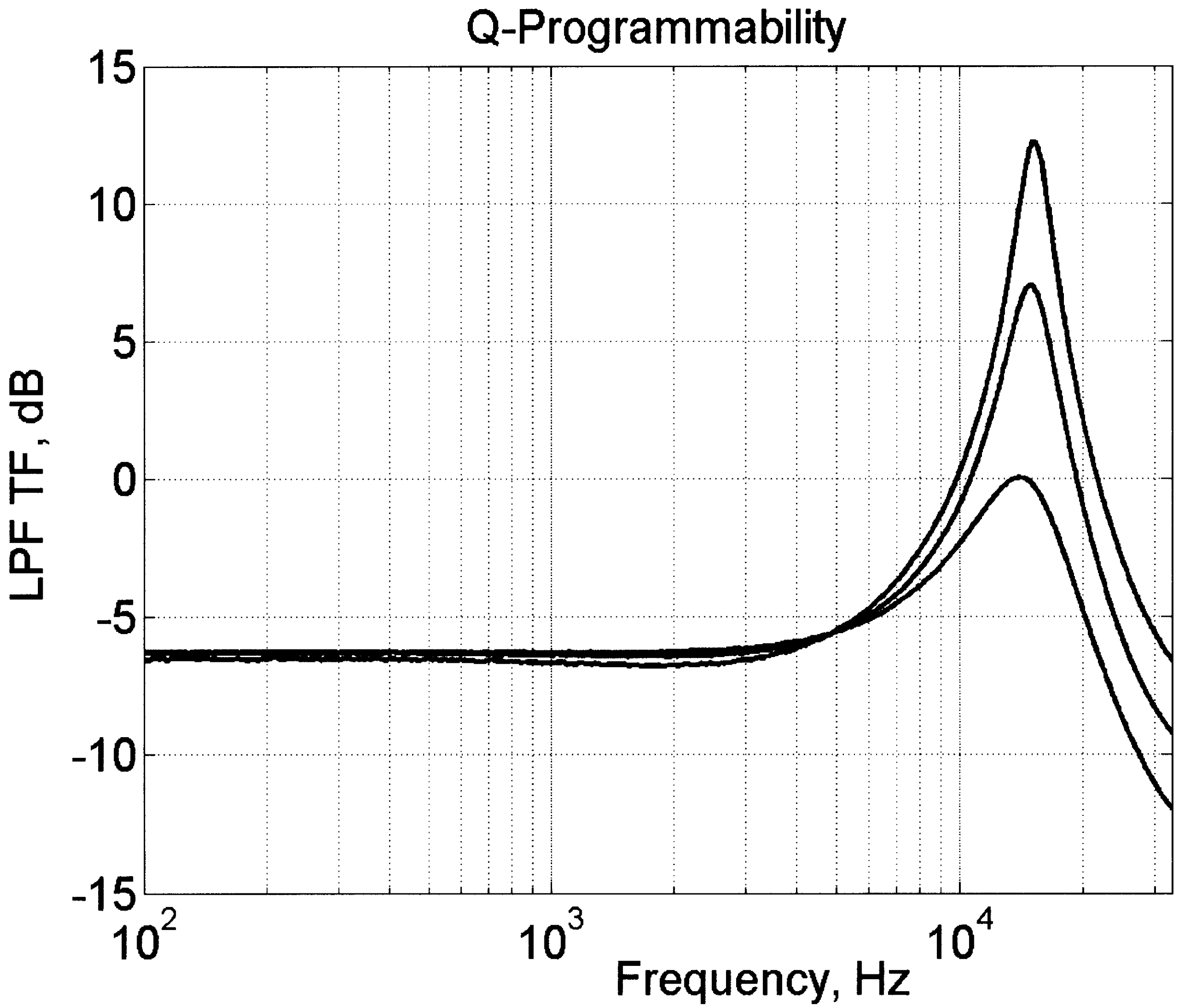


Figure 9: Transfer functions for various quality factor settings.

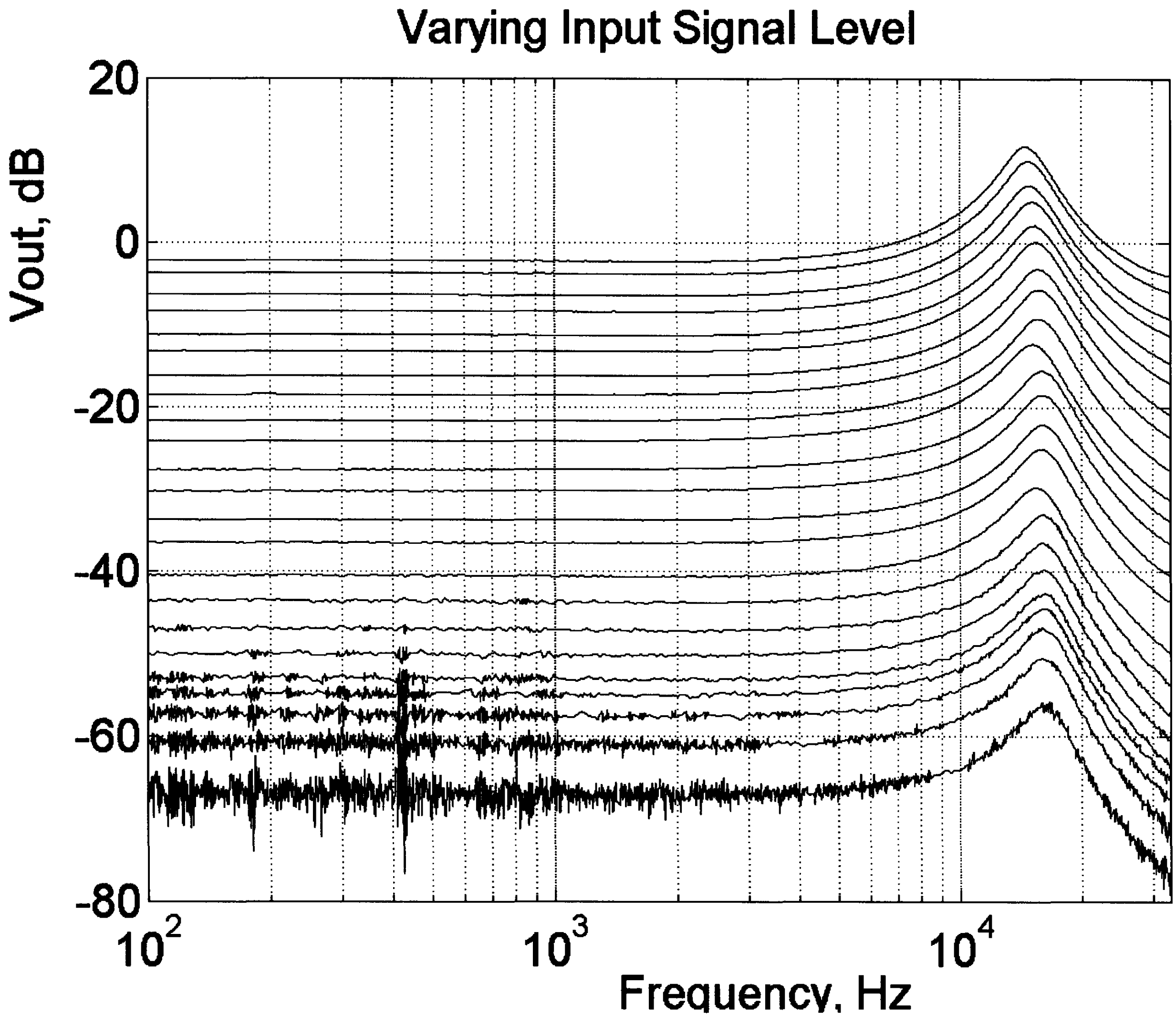


Figure 10: Output signal dependence on frequency is shown for the input signal magnitude varying over 70 dB range.

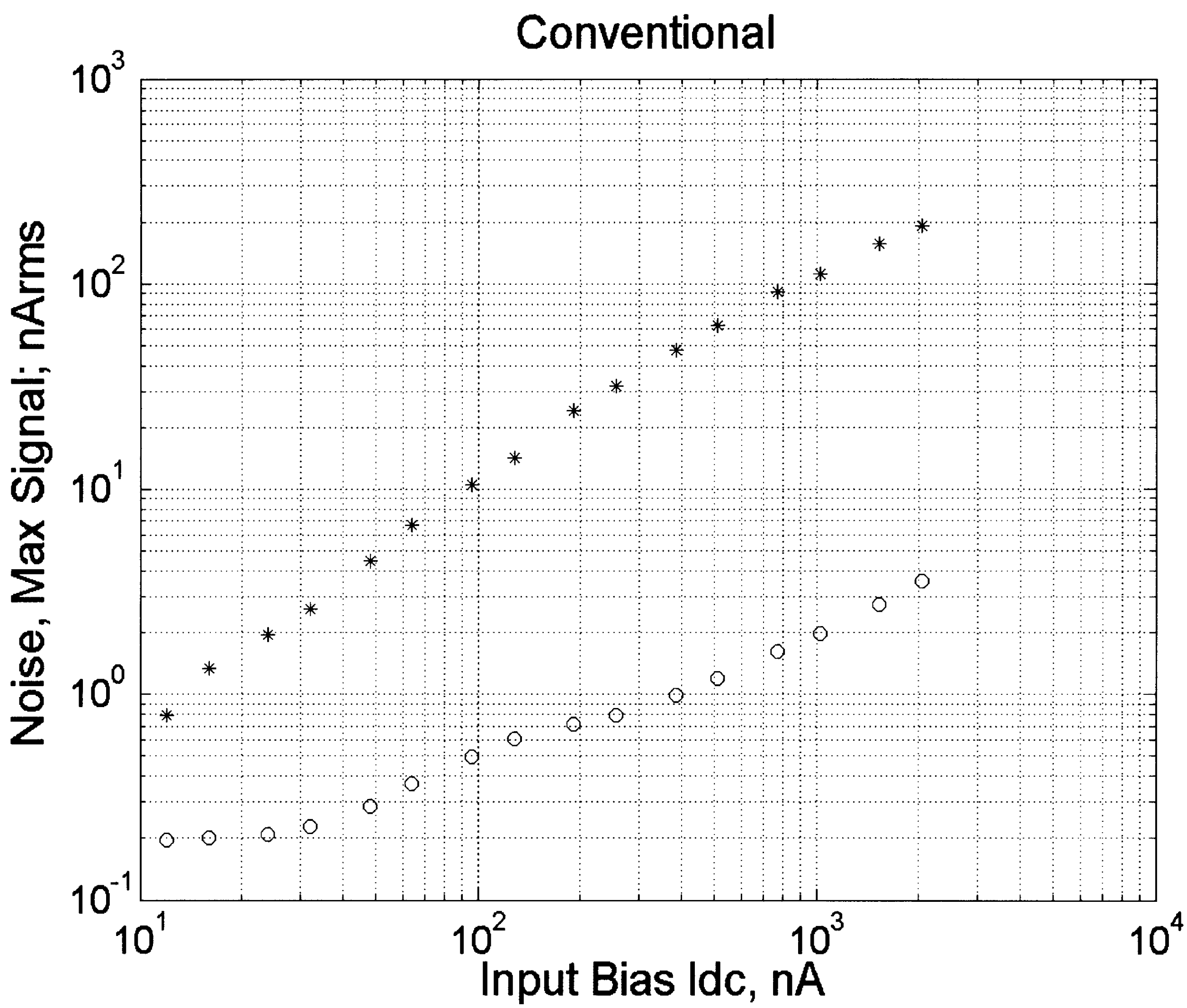


Figure 21: Maximum output signal (*) and noise at the output (o) dependence on the input bias current for the conventional log-domain design approach of Figure 6. The ratio of the input signal amplitude to the input DC bias was adjusted to hold the THD approximately constant.

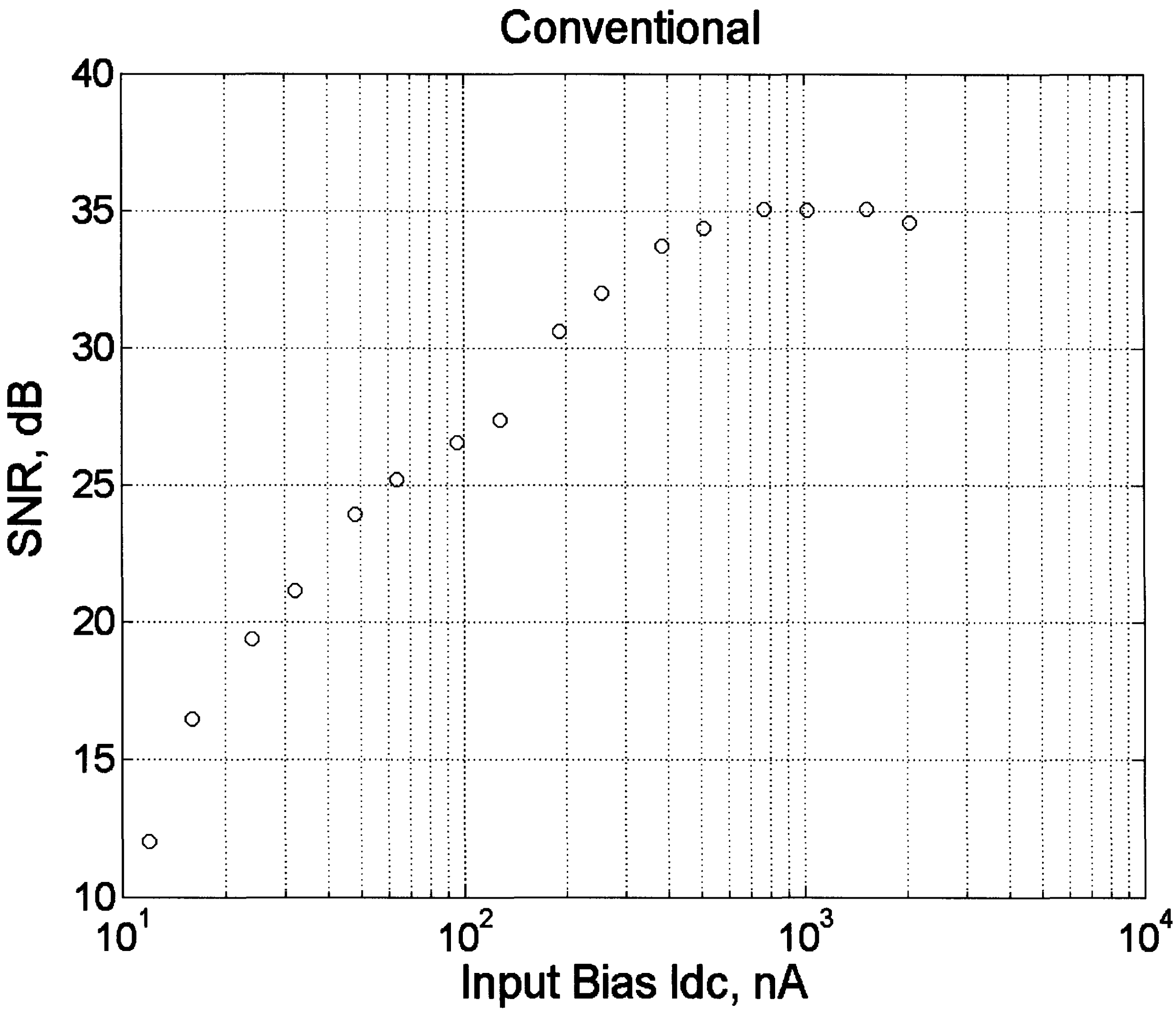


Figure 12: SNR versus the input DC bias for the conventional log-domain design. The SNR was computed from Figure 11. Figures 11 and 12 demonstrate the maximum SNR of 35.1 dB and the DR of 56 dB. Quality factor Q=4.

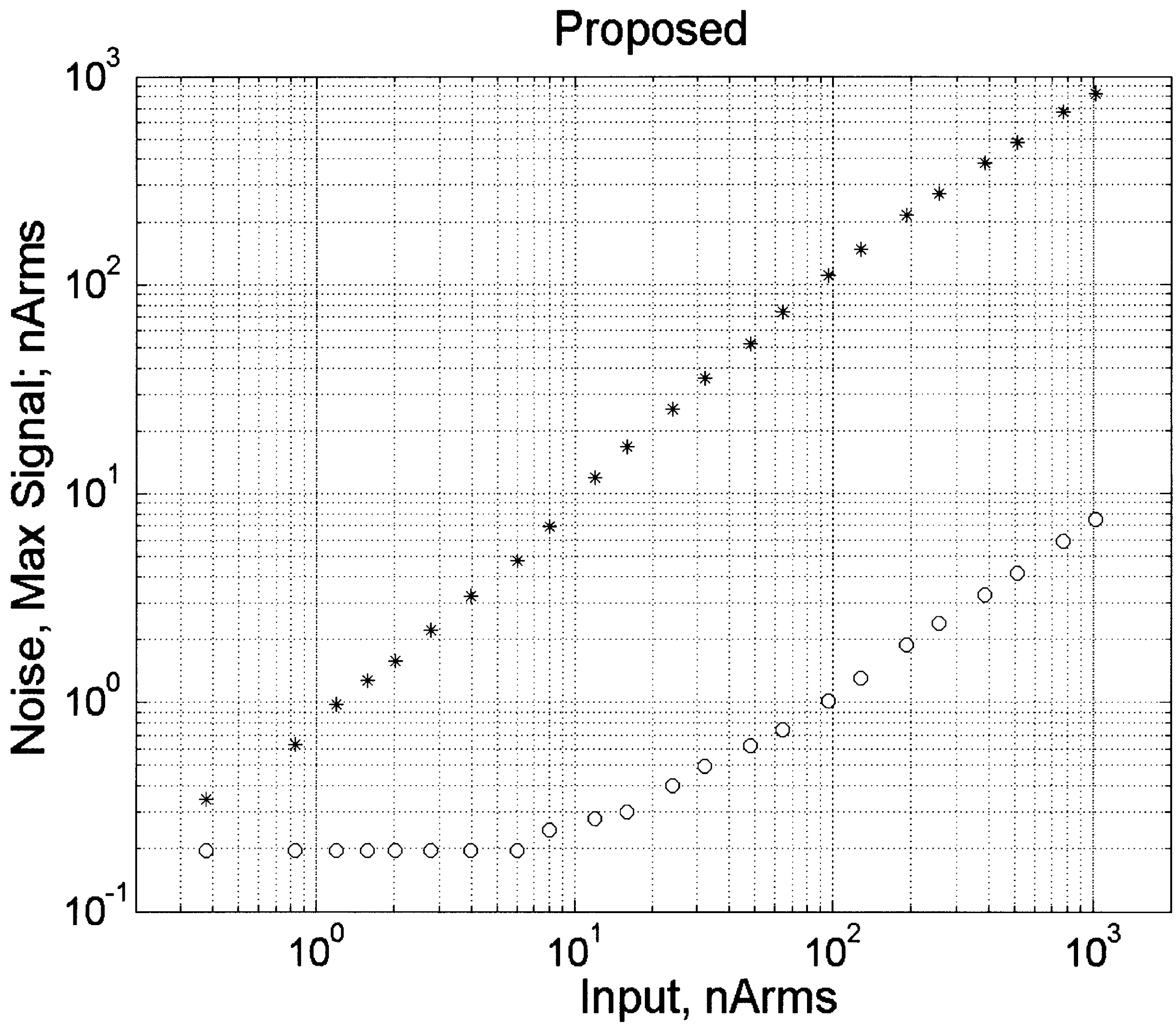


Figure 13: Maximum output signal (*) and noise at the output (o) as a function of the input signal level for the proposed method. The ratio of the input signal amplitude to the input bias was 0.8, which resulted in approximately steady THD equal to that observed in the measurements of Figures 11 and 12.

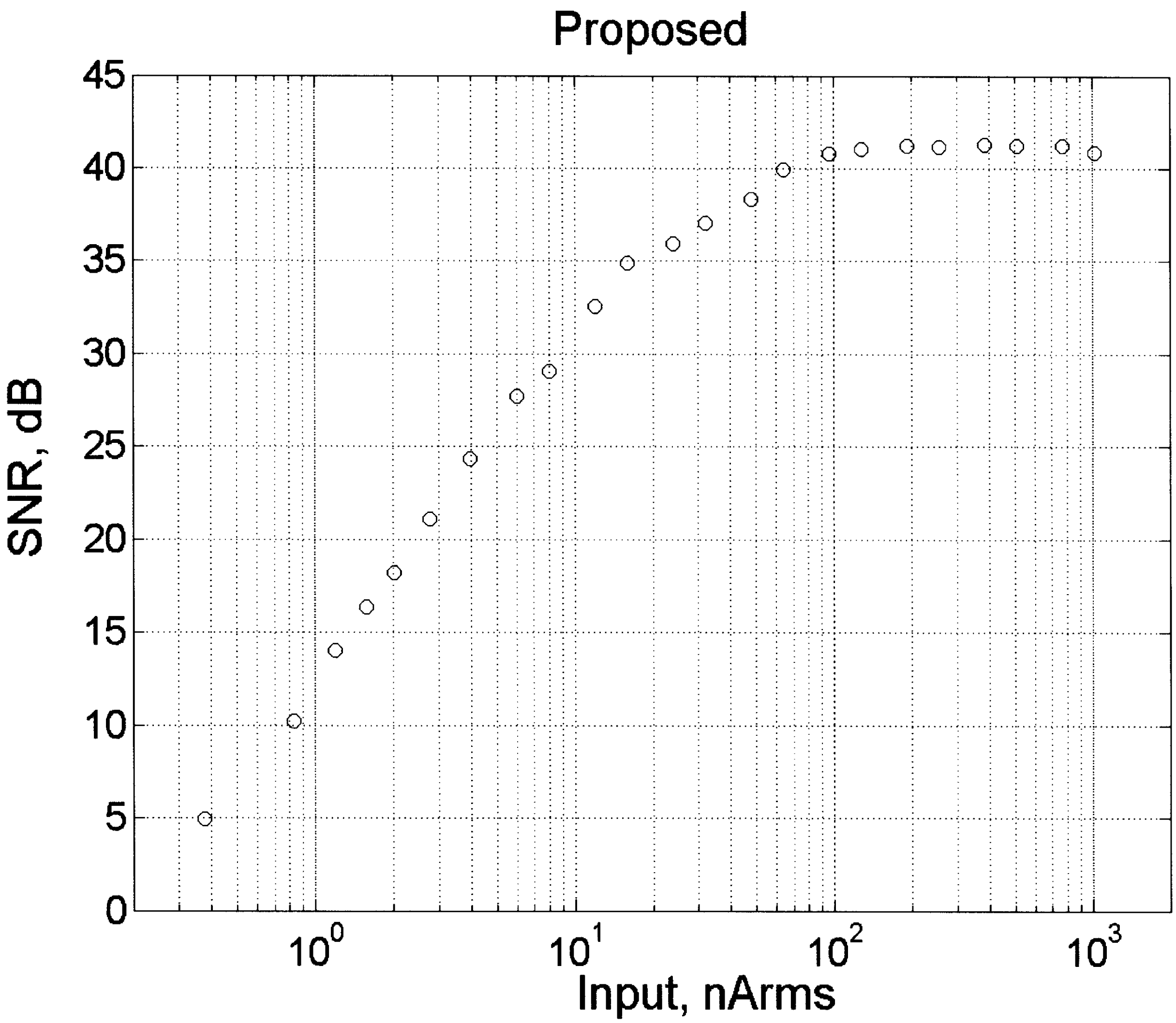


Figure 14: SNR dependence on the input signal level for the proposed log-domain design technique. The SNR was calculated from Figure 13. Figures 13 and 14 demonstrate the maximum SNR of 41.3 dB and the DR of 76 dB. Quality factor Q=4.

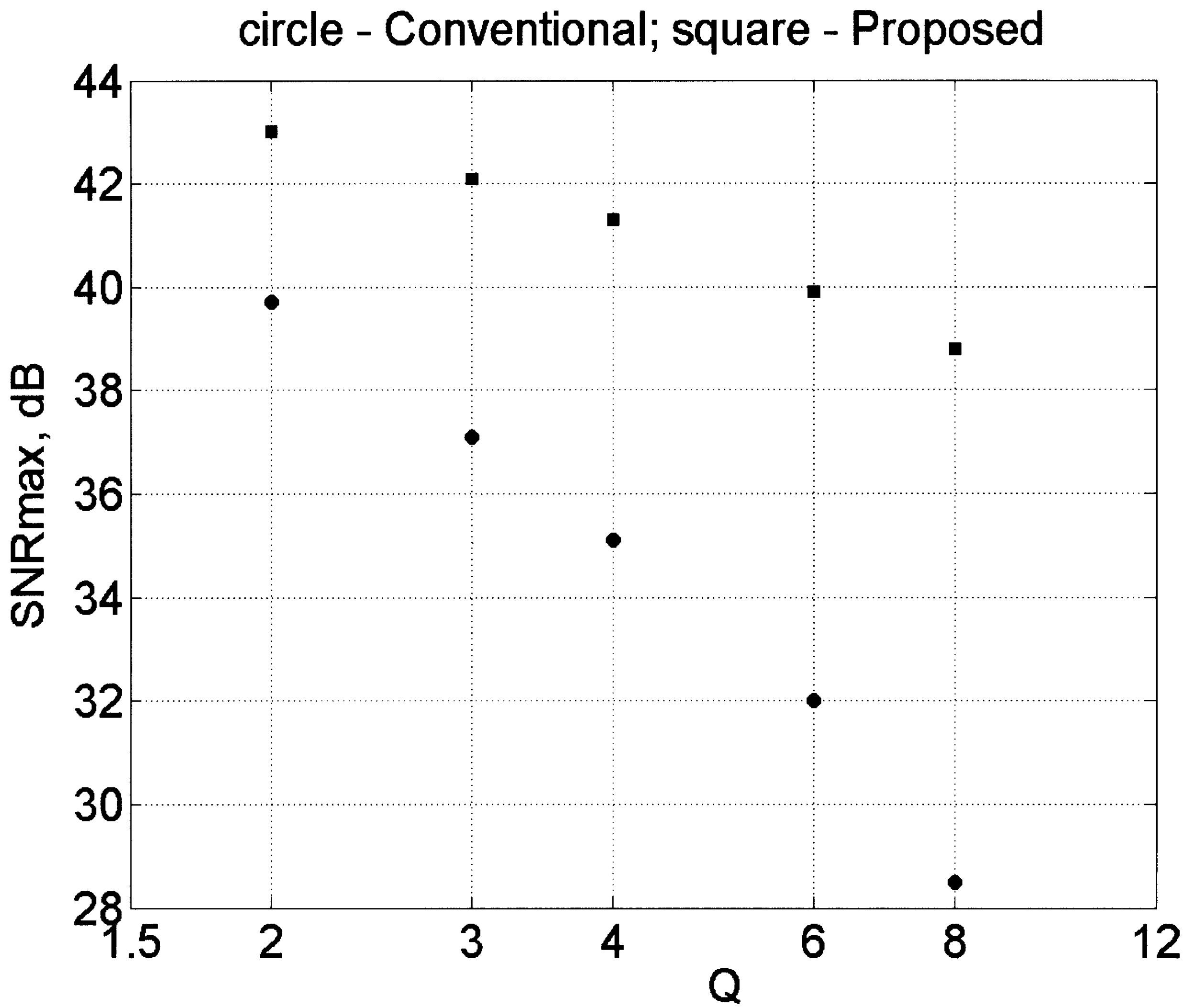


Figure 15: As the filter is programmed for higher Q, the maximum achievable SNR degrades. In the conventional filter of Figure 6 this degradation (shown in circles) is proportional to Q^2 in terms of signal power, whereas in the proposed filter the degradation (shown in squares) is proportional to only Q.

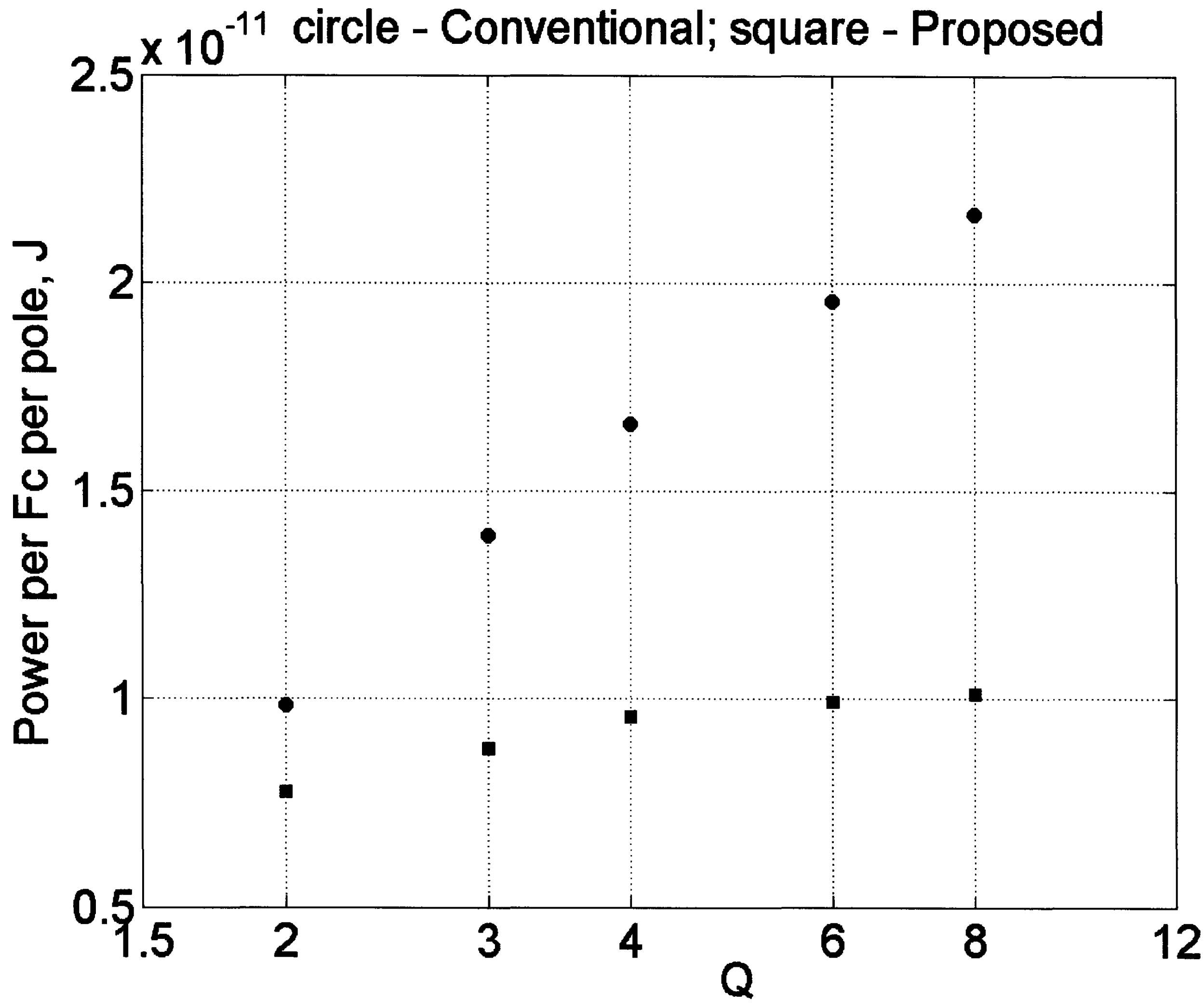


Figure 16: In the conventional log-domain design the power consumption (shown in circles) normalized to corner frequency and the number of poles (Figure of Merit) degrades as the filter is programmed for higher Q. In the proposed filter the Figure of Merit (FOM) stays approximately constant (shown in squares).

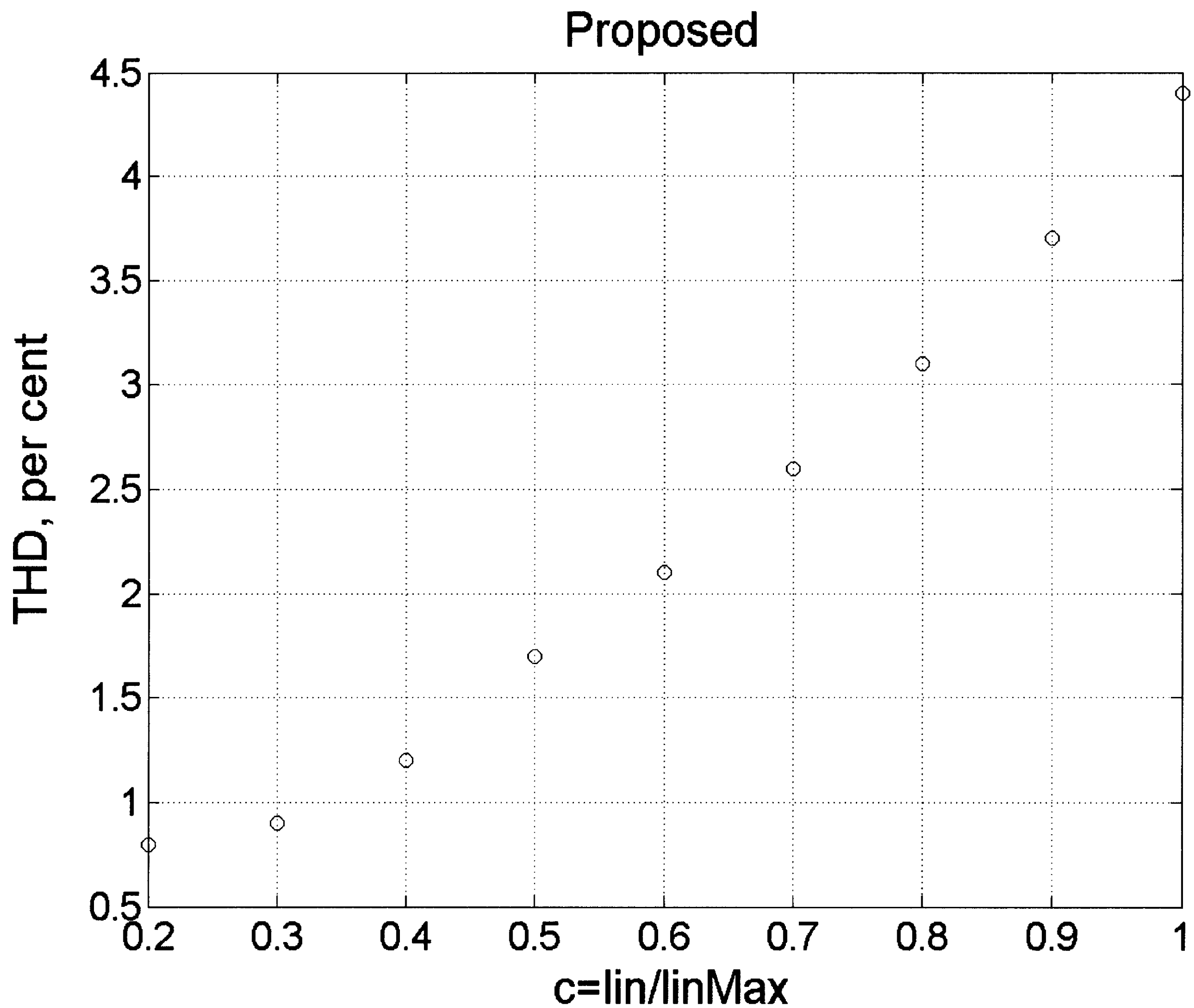


Figure 17: Harmonic distortion of the output signal in the proposed filter as a function of the ratio of the input signal amplitude to the input bias measured at the corner frequency of the filter.

6. Electronic Cochlea

Abstract - Silicon cochleas are inspired by the biological cochlea and perform efficient spectrum analysis: They realize a bank of constant-Q N^{th}-order filters with O(N) efficiency rather than $O(N^2)$ efficiency due to their use of an exponentially tapered filter cascade. They are useful in speech-recognition front ends, cochlear implants, and hearing aids, especially as architectures for improving spectral analysis in noisy environments and for performing low-power spectrum analysis. In this Chapter we describe a current-mode 33-stage 0.18-μm silicon cochlea that achieves 79-dB of dynamic range with 41-μW power consumption on a 1-V power supply over a usable 3.5kHz-14kHz frequency range. These numbers represent an 18dB improvement in dynamic range and a 12.5x reduction in power consumption over prior state-of-the-art silicon cochleas.

6.1. Ultra-low-power wide-dynamic-range front-end

In this section we describe the ultra-low-power wide-dynamic-range front-end of the cochlea chip, which is able to accept either current inputs from MEMS or commercial electret microphones, or auxiliary voltage inputs.

The most straightforward way to input a signal from an off-chip current source onto a chip is shown in Figure 1 (A).

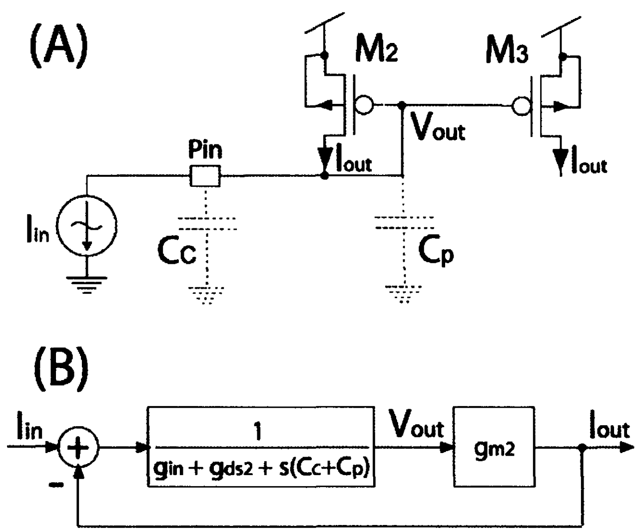


Figure 1: (A) The simplest way to input an off-chip current signal onto a chip; the current I_{in} has to be always positive, which can be achieved through class-A technique by adding to the AC signal the DC current that can be adjusted as the signal level varies. (B) Block diagram showing the first-order low-pass transfer function of this circuit with bandwidth $g_{m2}/(C_c+C_p)$ limited by the sum of the off-chip and on-chip parasitic capacitances C_c and C_p. Here g_{in} is the current source's I_{in} conductance, g_{m2} and g_{ds2} are the transconductance and the output conductance of M2.

This circuit is a simple current mirror; it is the first-order low-pass passive RC-filter with the bandwidth determined by the ratio of the incremental conductance g_{m2} of the diode-connected transistor M2 to the sum of the off-chip C_c and on-chip C_p parasitic capacitances. We can also treat this circuit as a negative feedback system. The block diagram of such feedback system is shown in Figure 1 (B). The loop gain has one left-half-plane pole on the real axis, so the system is always stable. To compute the unity-gain cross-over frequency that is equal to the bandwidth for the first-order systems, we ignore the conductance g_{in} of the current source and the output conductance g_{ds2} of M2 to obtain:

$$\omega_1 = g_{m2} / \left(C_C + C_p \right) \tag{1}$$

Here g_{m2} is the transconductance of M2. In the subthreshold region of CMOS operation:

$$g_{m2} = \kappa_p \cdot I_{in,DC} / U_t \tag{2}$$

Here κ_p is the subthreshold exponential parameter of the PMOS transistors, U_t is the thermal voltage kT/q, and $I_{in,DC}$ is the DC component of the input signal I_{in}. The drain current I_{out} has to be always unidirectional; therefore, I_{in} has to be unidirectional, which can be achieved through the class-A technique by adding the DC current that is equal or exceeds the maximum amplitude of the signal I_{AC}. However, this method is power-hungry and adds a lot of noise to the output I_{out}. We would like to adjust $I_{in,DC}$ as the signal level varies to follow the signal's I_{AC} envelope. However, for soft signals the transconductance g_{m2} becomes small, and (1) shows that the front-end becomes slow. The fast-alternating input current I_{AC} would not propagate to the output, being used up to change the voltage on the parasitic capacitance C_c+C_p. We add an amplifier into the feedback path to reduce the voltage swing on the pin and mitigate the effect of the off-chip parasitic capacitance C_c. Figure 2 (A) shows the front-end circuit with added common-gate amplifier M1. Figure 2 (B) shows the block diagram of this feedback system. The leftmost block in this diagram contains the additional term g_{m1}, the transconductance of M1, in the denominator. This is because the node V_{gnd} sees the input conductance of the common-gate amplifier M1 that is approximately equal to the transconductance of M1, g_{m1}. However, near the unity-gain cross-over frequency the $s{\cdot}C_c$ term still dominates the sum of the conductances so that we can ignore the latter.

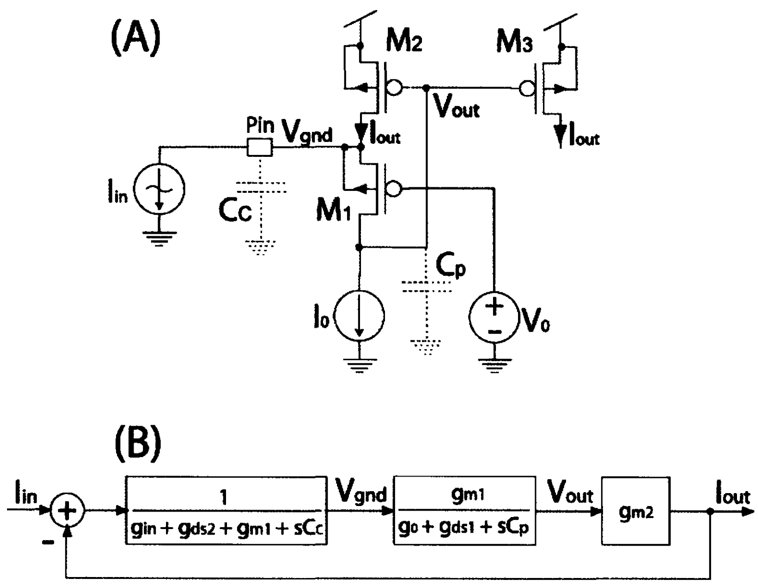


Figure 2: (A) Common-gate amplifier M1 is added into the feedback path to reduce the voltage swing V_{gnd} on the pin and the off-chip capacitance C_c. (B) Block diagram indicating the speed-up of the circuit by a factor of $A_1 \approx g_{m1}/(g_0+g_{ds1}) \approx \kappa_p \cdot V_E/(2 \cdot U_t)$. Here A_1 is the voltage gain of the common-gate amplifier, g_{m1} and g_{ds1} are the transconductance and output conductance of M1, g_0 is the conductance of the current mirror I_0 approximately equal to g_{ds1}, κ_p is the subthreshold exponential parameter of the PMOS transistors, V_E is the Early voltage, and U_t is the thermal voltage kT/q. However, the on-chip parasitic capacitance C_p can cause stability problems; then, C_c needs to be increased, but the speed of the circuit is determined by the on-chip parasitic C_p as opposed to much larger off-chip C_c.

The voltage gain of the common-gate amplifier M1 is represented by the second block in the diagram, where g_0 is the conductance of the current mirror I_0 that is approximately equal to g_{ds1}, the output conductance of M1. In the first approximation, if we ignore the on-chip parasitic capacitance term $s \cdot C_p$, the unity-gain cross-over frequency is:

$$\omega_2 = A_1 \cdot g_{m2} / C_C \approx A_1 \cdot \omega_1 \tag{3}$$

Where A_1 is the DC gain of the common-gate amplifier M1:

$$A_1 \equiv \frac{g_{m1}}{g_0 + g_{ds1}} \approx \frac{\kappa_p \cdot V_E}{2 \cdot U_t} \tag{4}$$

Here V_E is the Early voltage. The equation (3) demonstrates the increase in bandwidth by the factor A_1 – at negligible cost in power consumption and noise, if the current I_0 is chosen near the amplitude of the softest expected signal I_{AC}, i.e., near the minimum value of $I_{in,DC}$. However, taking into account the on-chip parasitic capacitance C_p, we note that the system response is second-order. To avoid the overshoot and ringing, we need to ensure that the phase margin of our system is acceptable. If we require the phase margin to be greater than 45°, for example, then C_p should be small enough to guarantee:

$$\omega_3 \equiv \frac{g_0 + g_{ds1}}{C_p} \geq \omega_2 = \frac{A_1 \cdot g_{m2}}{C_C} \tag{5}$$

The above equation can be rewritten to explicitly illustrate the so-called A^2-stability problem that circuits with two high-gain nodes have:

$$\frac{C_P}{C_C} \leq \frac{1}{A_1^2} \cdot \frac{g_{m1}}{g_{m2}} = \frac{1}{A_1^2} \cdot \frac{I_0}{I_{in,DC}} \tag{6}$$

In the 0.18-µm process that was used to fabricate the electronic cochlea the equation (6) held even for the maximum value of $I_{in,DC}$. But for older processes, where C_p is larger relative to the off-chip parasitic capacitance C_c and the equation (6) does not hold, the additional explicit compensating capacitor would need to be used so that the response of the front-end is acceptable. In this case the bandwidth is determined by ω_3 and the speed-up over the simple current mirror is smaller than A_1. Somewhat better results could be achieved if the resistor is added in series to the external compensating capacitor to form an off-chip lead compensation network.

Implementation of the adjustable current $I_{in,DC}$ that follows the envelope of the signal I_{AC} requires an envelope detector. Chapter 2 describes the envelope detector comprising the rectifier shown in Figure 3 (A) and the peak detector with asymmetric attack and release time constants shown in Figure 3 (B).

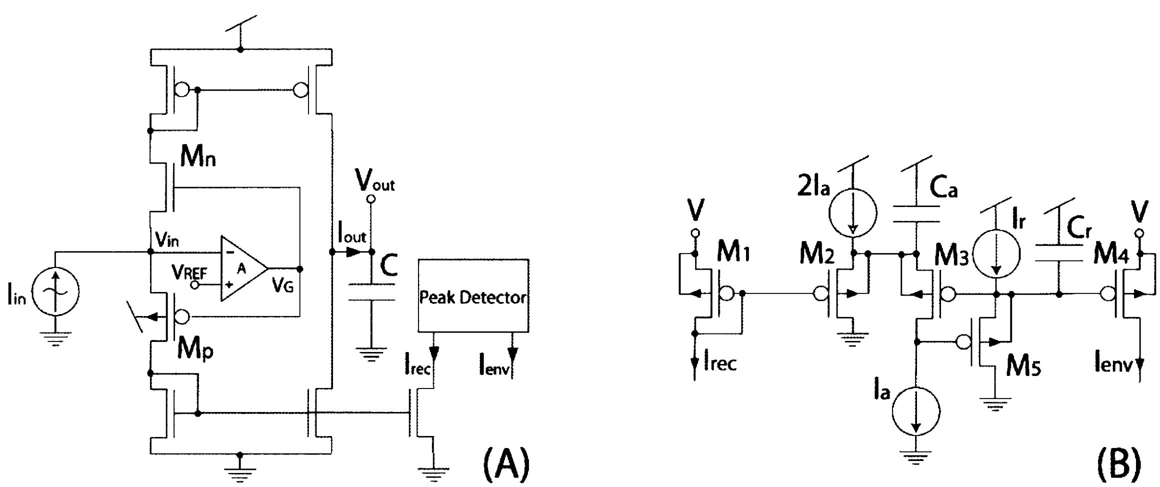


Figure 3: (A) Rectifier from Chapter 2; I_{out}=-I_{in} within the dynamic range of the operation, I_{rec} is either half-wave rectified (shown here) or full-wave rectified signal, and I_{env} is the envelope of the signal. (B) Peak detector circuit from Chapter 2 with the asymmetric attack and release time constants.

Figure 4 shows the proposed ultra-low-power wide-dynamic-range front-end circuit. The minor negative feedback loop comprising the rectifier, the transconductor M7 and the sped-up current mirror M1-M3 eliminates the low-frequency components of the current I_{AC} flowing into the V_2 and V_1 nodes by injecting the current into the V_{gnd} node such that the low-frequency components of the drain current of M3 become equal to those of M5. Therefore, the minor loop ensures the rectifier's normal operation. The major negative feedback loop additionally includes the peak detector and the current mirrors M4-M6. This loop provides the bias current I_{env} that adjusts as the I_{AC} amplitude varies ensuring the normal operation of the entire system. The analysis of the major loop is complicated, but it is obvious that fast attack time of the peak detector is crucial for its stability.

The major loop adapts the bias currents to the signal level, reducing the quiescent power consumption and noise. This adaptive biasing increases the SNR for low signal levels extending the dynamic range of the system by the dynamic range of the rectifier. The proposed circuit can accept current inputs from MEMS or commercial electret microphones, or auxiliary voltage input V_{in} as shown in Figure 4, or even all of the above simultaneously. This front-end has very low distortion in converting voltage signals into the current as this conversion happens in the linear resistor R connected to the virtual ground node V_{gnd}. The proposed front-end is useful for any current-mode circuit, but adaptive bias topologies benefit the most.

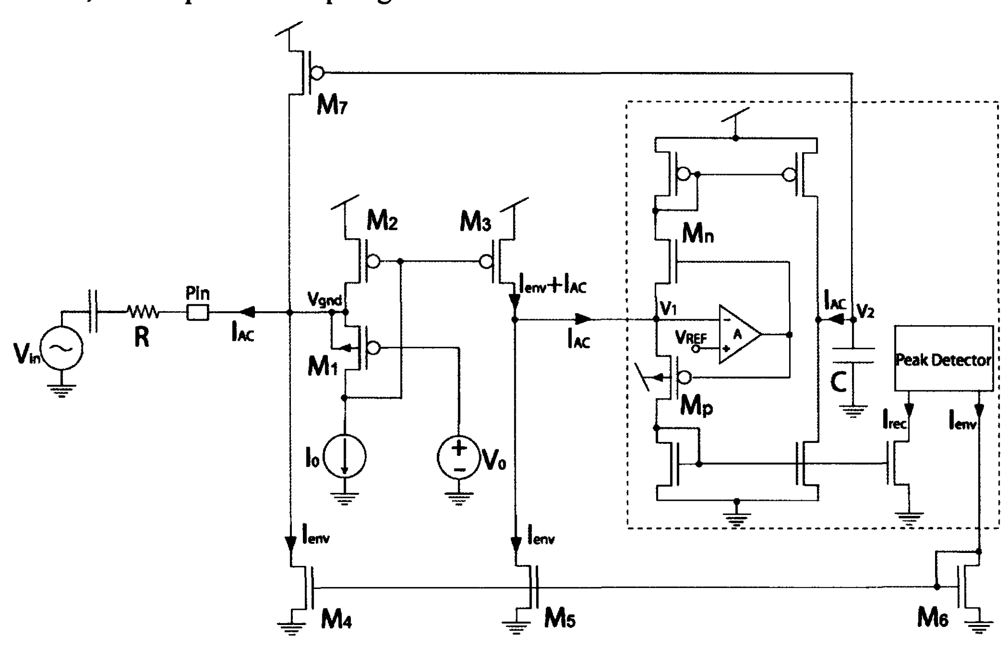


Figure 4: Ultra-low-power wide-dynamic-range front-end of the cochlea chip, which is able to accept either current inputs from MEMS or commercial electret microphones, or auxiliary voltage input V_{in} as shown. The output current is mirrored out from M3, the rectified and envelop detected versions are also available. When the input signal is soft, the currents are low contributing little to noise and power consumption. When the signal is large, the envelop detector provides the necessary current in the feedback loop to avoid distortion.

6.2. Stage's All-Pass Filter Implementation

6.2.1. Practical All-Pass Transfer Function

In Chapter 3 we designed the transfer function for the stage of the cochlear cascade to be (see (31) in 3.4.1):

$$H(s)=\frac{s^2+2d\cdot s+1}{(1+\alpha)s^2+(2d+\alpha\cdot\mu)\cdot s+1} \tag{7}$$

Here $s=j\cdot\omega$, ω is the frequency normalized to the spatial-dependent corner frequency, d is the signal level dependant damping factor that realizes the distributed gain control in the cochlear cascade, α and μ are the constants.

But the exact implementation of this transfer function is impractical. To arrive at the transfer function that permits efficient implementation and maintains the properties needed to realize the cochlear cascade that we discussed in Chapter 3, we consider transfer functions in the following form:

$$H(s)=k_{in}+k_1\cdot\frac{1}{(1+\alpha)s^2+(2d+\alpha\cdot\mu)\cdot s+1} \tag{8}$$

We choose the coefficients k_{in} and k_1 such that the group delays of the transfer functions in (7) and (8) are the same, and also $k_{in}+k_1=1$ to maintain unity low-frequency gain. These conditions give:

$$k_1=\frac{\alpha\cdot\mu}{2d+\alpha\cdot\mu} \qquad k_{in}=\frac{2d}{2d+\alpha\cdot\mu} \tag{9}$$

The coefficients in (9) are signal level dependant, but the transfer function described by (8) with (9) can be efficiently implemented in analog log-domain circuits. Figure 8 (A) shows the magnitude of such a transfer function with $\alpha=0.7$, $\mu=0.2$, and $2d$ varying from 0.2 to 0.8. Cascading the stages with the above parameter values and exponentially tapered corner frequencies at 12 stages per octave realizes the cochlea with 44-dB of collective gain.

6.2.2. Circuit Realization of the All-Pass Filter

Figure 5 shows the log-domain implementation of the second-order low-pass filter adapted from Chapter 5 (see Figure 2 in Chapter 5). The compressed input signal is supplied to V_{in} terminal, scaled and compressed input signal's envelope is supplied to V_a terminal, and the voltage V_{out} is expanded into an output current.

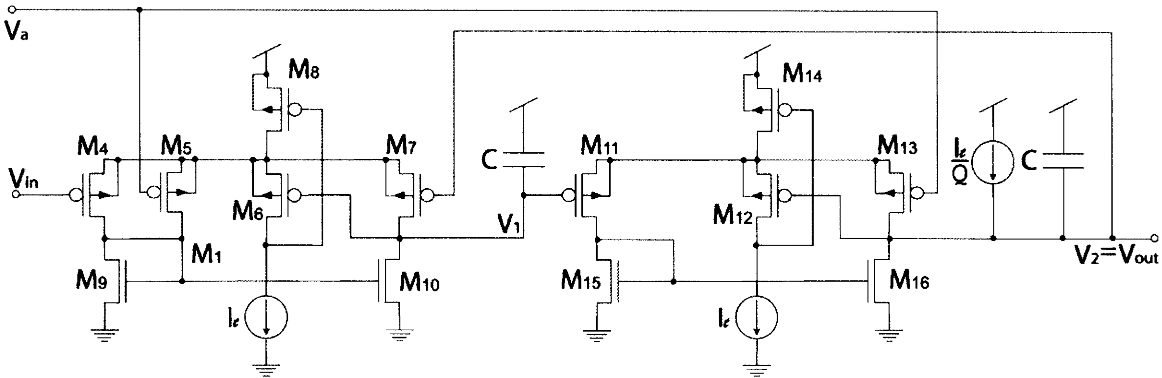


Figure 5: The low-pass filter (LPF) log-domain core from Chapter 5. The compressed input signal is supplied to V_{in} terminal, scaled and compressed input signal's envelope is supplied to V_a terminal, and the voltage V_{out} is expanded into an output current. The corner frequency of this filter is controlled by I_t, and the quality factor Q is programmed by I_t/Q.

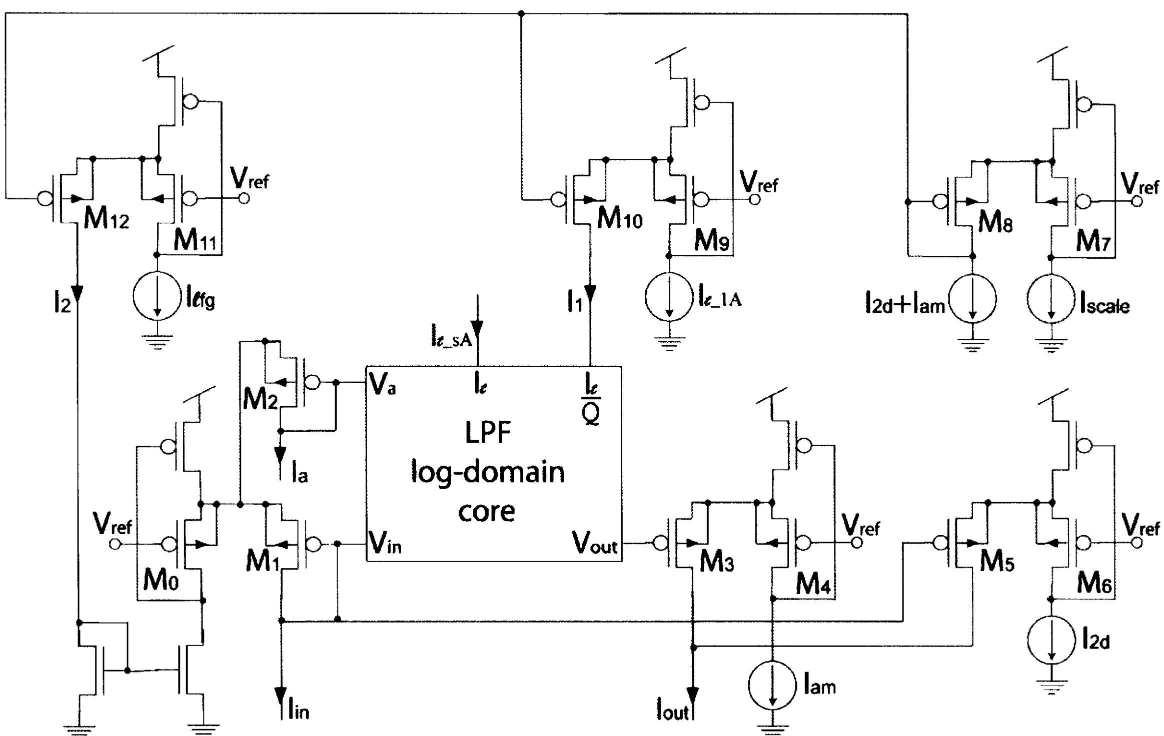


Figure 6: Implementation of the coefficients k_l and k_{in} from (9) to realize the complex zero pair and the gain control. Only one half of pseudo-differential system is shown.

The corner frequency of this filter is controlled by I_t, and the quality factor Q is programmed by I_t/Q. This log-domain core realizes the second term of the transfer function in (8). Figure 6 shows the implementation of the stage's filter. The circuit in Figure 6 realizes the following transfer function:

$$H(s)=\frac{I_{scale}}{I_{lfg}}\cdot\left(\frac{I_{2d}}{I_{2d}+I_{am}}+\frac{I_{am}}{I_{2d}+I_{am}}\cdot\frac{1}{1+s\cdot\frac{I_{t\_1A}\cdot(I_{2d}+I_{am})}{I_{t\_sA}^{2}\cdot I_{scale}}\cdot\frac{C\cdot U_t}{\kappa_p}+s^2\cdot\left(\frac{C\cdot U_t}{\kappa_p\cdot I_{t\_sA}}\right)^2}\right) \tag{10}$$

Here C is the capacitor value in the log-domain low-pass core, the current I_{2d} is the scaled version of the envelope of the output signal from the previous stage implementing the feed-forward gain control, the currents I_{t_sA} and I_{t_1A} decrease exponentially along the cochlear cascade to implement the exponential tapering of the corner frequencies, and the rest of the currents are constants determined by the values of parameters such as α and μ.

6.3. Offset Current Cancellation and Efficient Rectification Circuit

To prevent the accumulation of the offset currents in the cascade, we need to have a low-frequency feedback loop that injects offset canceling current, similar to what we've done in the front-end circuit.

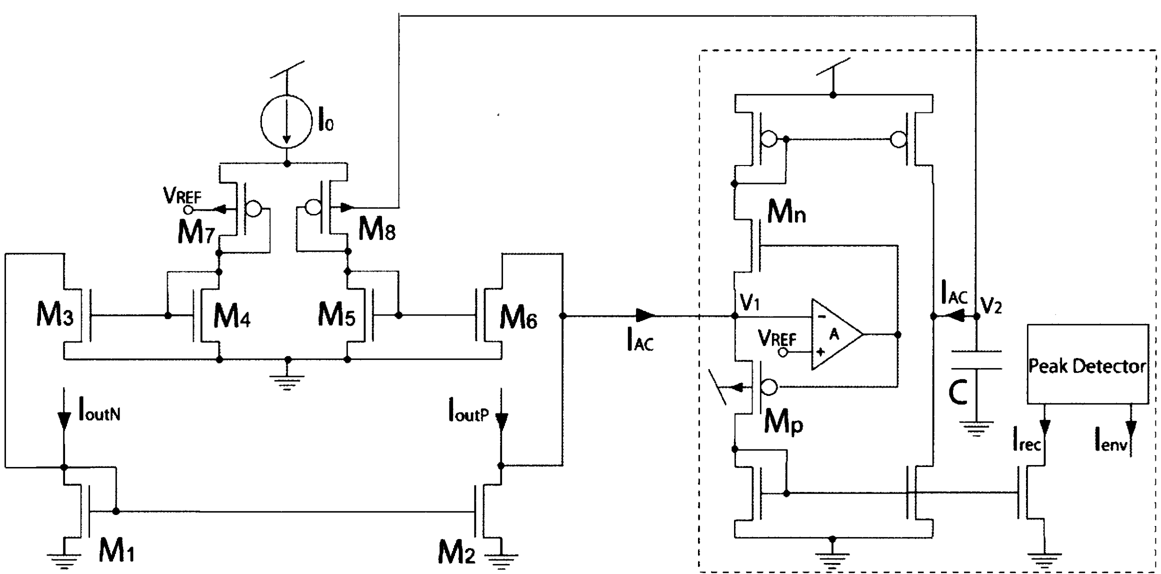


Figure 7: Offset current cancellation circuit. The feedback loop comprising the rectifier (in a dashed box) and the transconductor M3-M8 eliminates the low-frequency components of the current I_{AC} by injecting the offset canceling current into the difference of the pseudo-differential filter output currents I_{outN} and I_{outP}. This circuit also provides the half-wave rectified current I_{rec} and the output's envelope I_{env}.

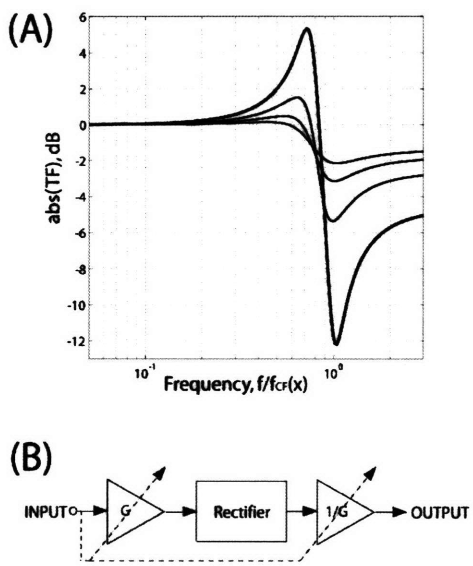


Figure 8: (A) Transfer function of the cochlear stage, which shows that the difference in output signal magnitudes of any consecutive stages does not exceed 18dB, implying that the envelope detection with 18dB dynamic range is sufficient to realize electronic cochlea's with any dynamic range. (B) Feedforward implementation of the efficient rectification in the cochlear cascade of all-pass filters, where the variable gain G is inversely proportional to the envelope of the previous stage's output signal.

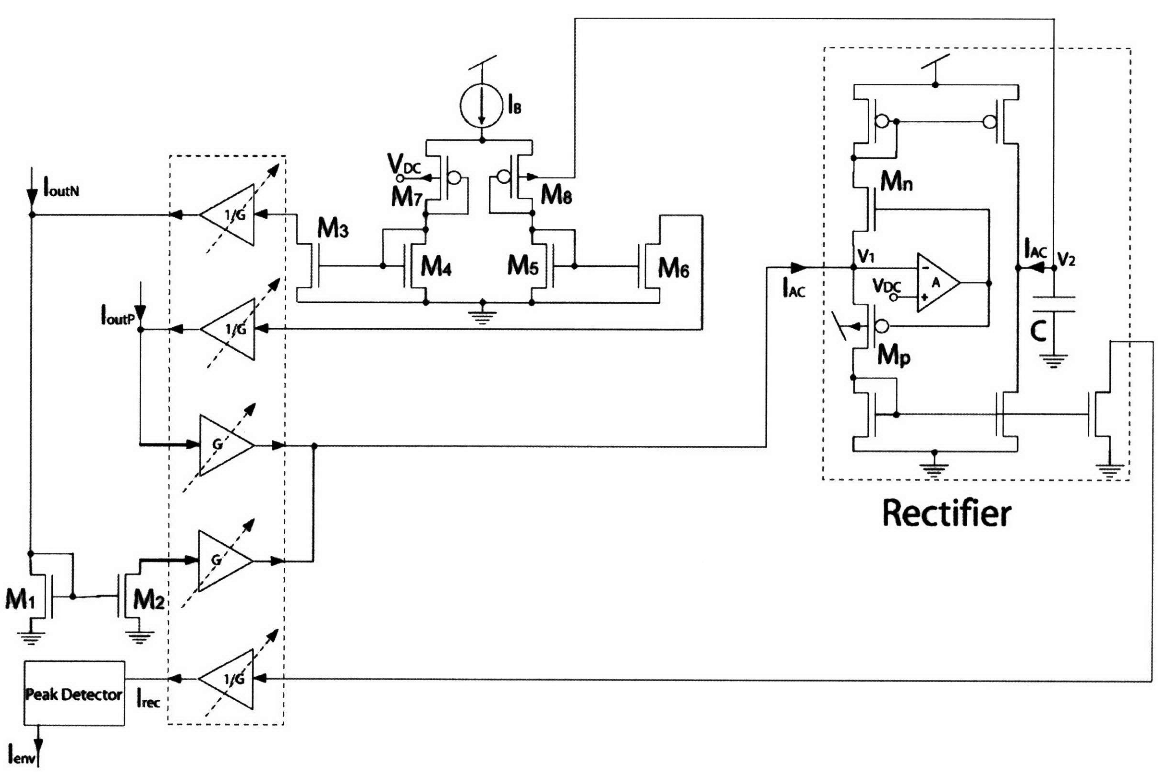


Figure 9: The efficient rectification from Figure 8 is coupled with adjustable offset cancellation. The maximum offset that can be cancelled is no longer limited by the bias current I_B of the transconductor M3-M8, but adjusts as $1/G$ proportionally to the signal level. This adjustable offset cancellation circuit allows reducing I_B and cutting its noise contribution, power consumption and capacitance C.

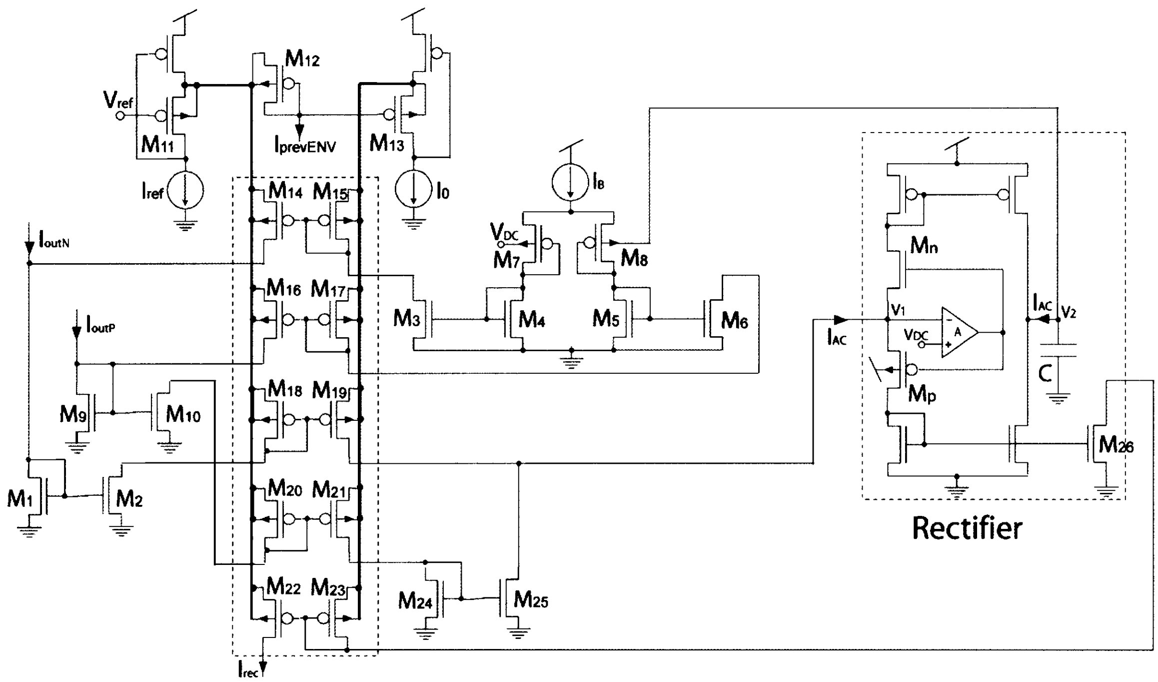


Figure 10: Transistor-level implementation of the efficient rectification and adjustable offset cancellation techniques. Transistors M12-M23 realize the variable gain amplifiers and attenuators G and $1/G$ controlled by the envelope of the previous stage's output $I_{prevENV}$. The offset-cancelled output signals that go to the next stage are obtained by subtracting the copies of the currents from transistors M2 and M10. The rectified signal I_{rec} goes to the peak detector circuit to obtain the output signal's envelope.

6.4. Experimental Results

6.4.1. 33-Stage Cochlea Chip

The experimental results from the chip:

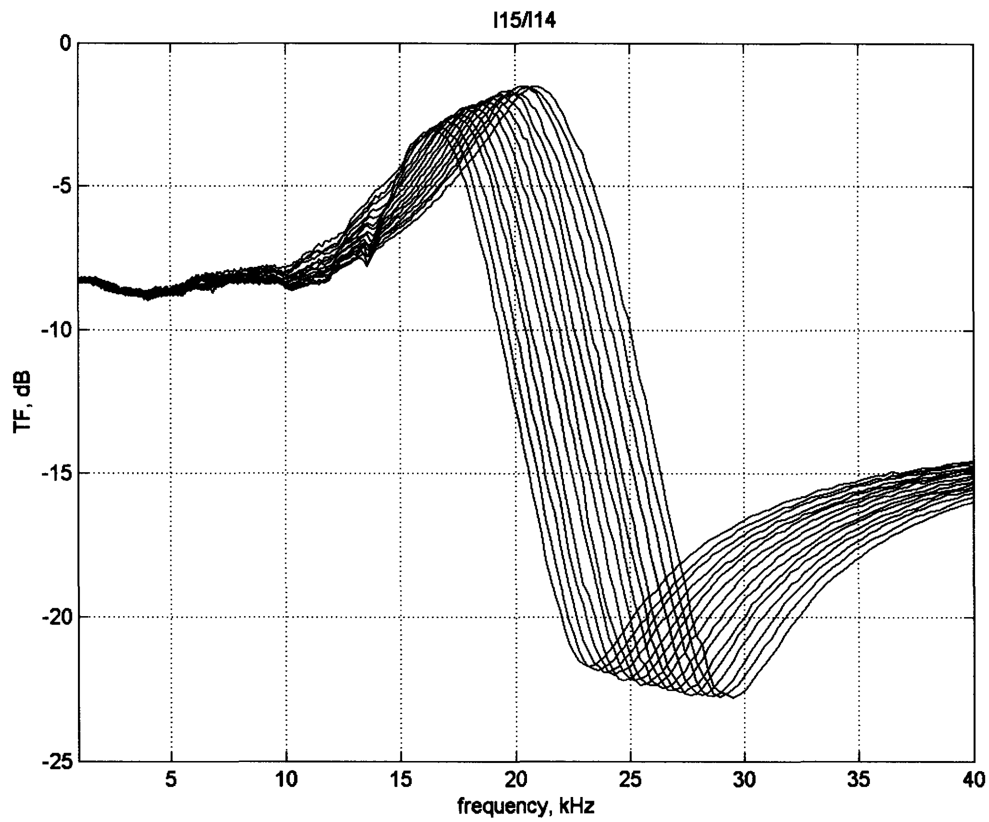


Figure 11: Transfer functions of the first stage are measured as the ratio of the envelope detector outputs of the first stage and the front-end. The corner frequency programming over the range of 16-kHz to 21-kHz is shown.

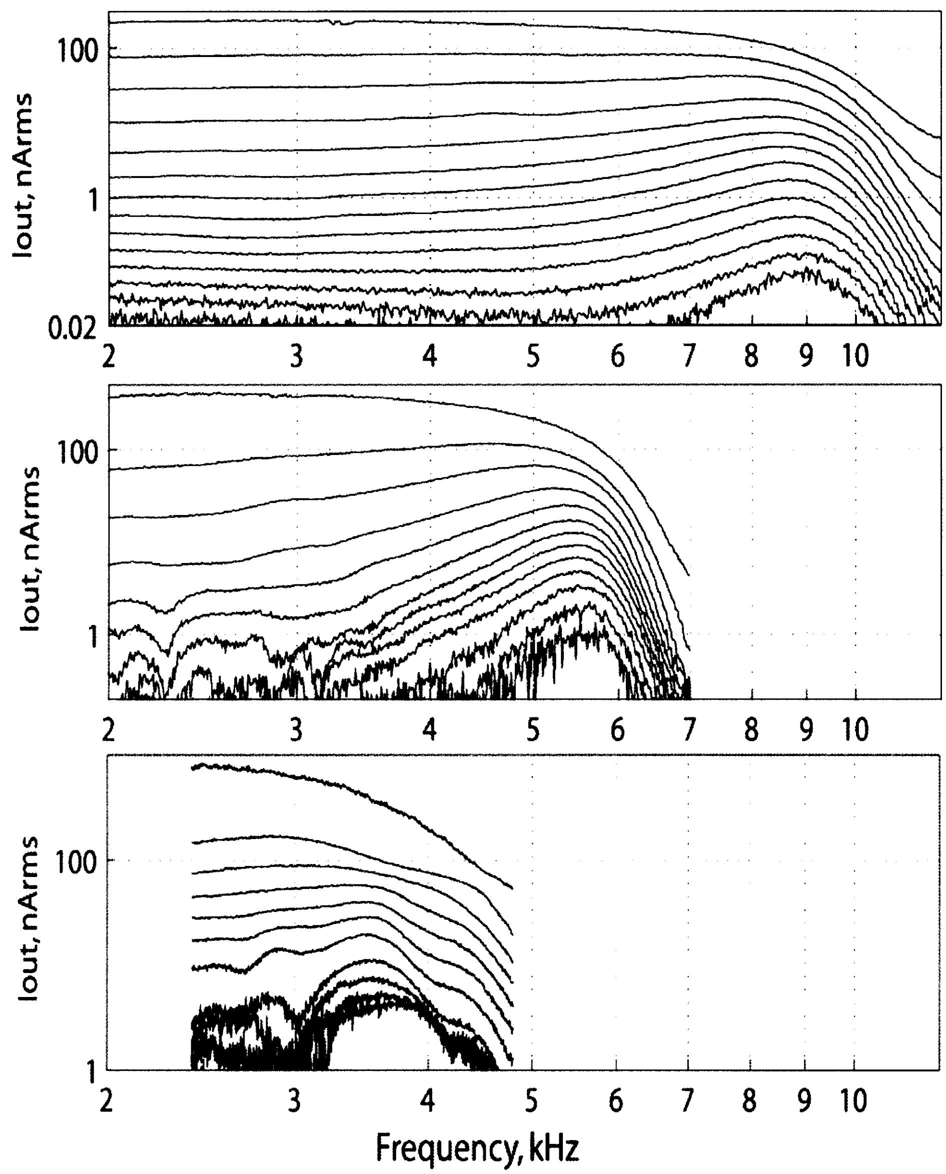


Figure 12: Magnitudes of frequency responses taken at outputs 9 (top), 17 and 23 (bottom) with varying intensity of the input signal are shown. This figure demonstrates the spatial distribution of the peak frequencies. The peak gain at low intensities is approximately 24-dB, and Q_{-10dB} is approximately 3.6. The usable frequency range is 3.5-14 kHz (best frequency range of 3.5-9 kHz is shown here).

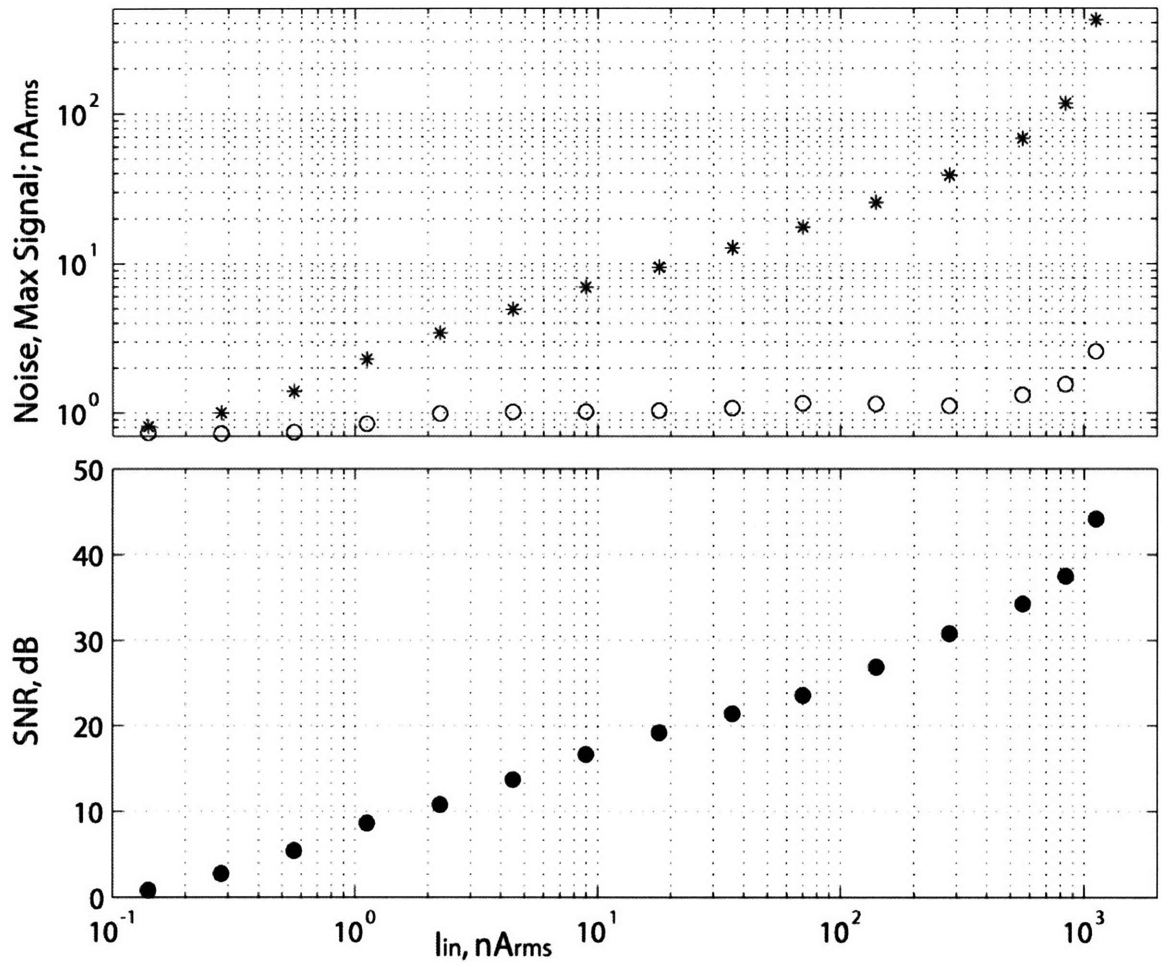


Figure 13: Top figure shows the noise (o) and the maximum signal (*) and the bottom figure shows SNR calculated from the top figure as a function of the input signal level. Output 17 is shown. As the input signal level increases, two conflicting effects take place: the increase in noise due to increasing bias current; and the decrease in the cochlear gain due to the gain control (compression) and the associated decrease in noise. This figure demonstrates the dynamic range of 79-dB and the maximum SNR of 44-dB.

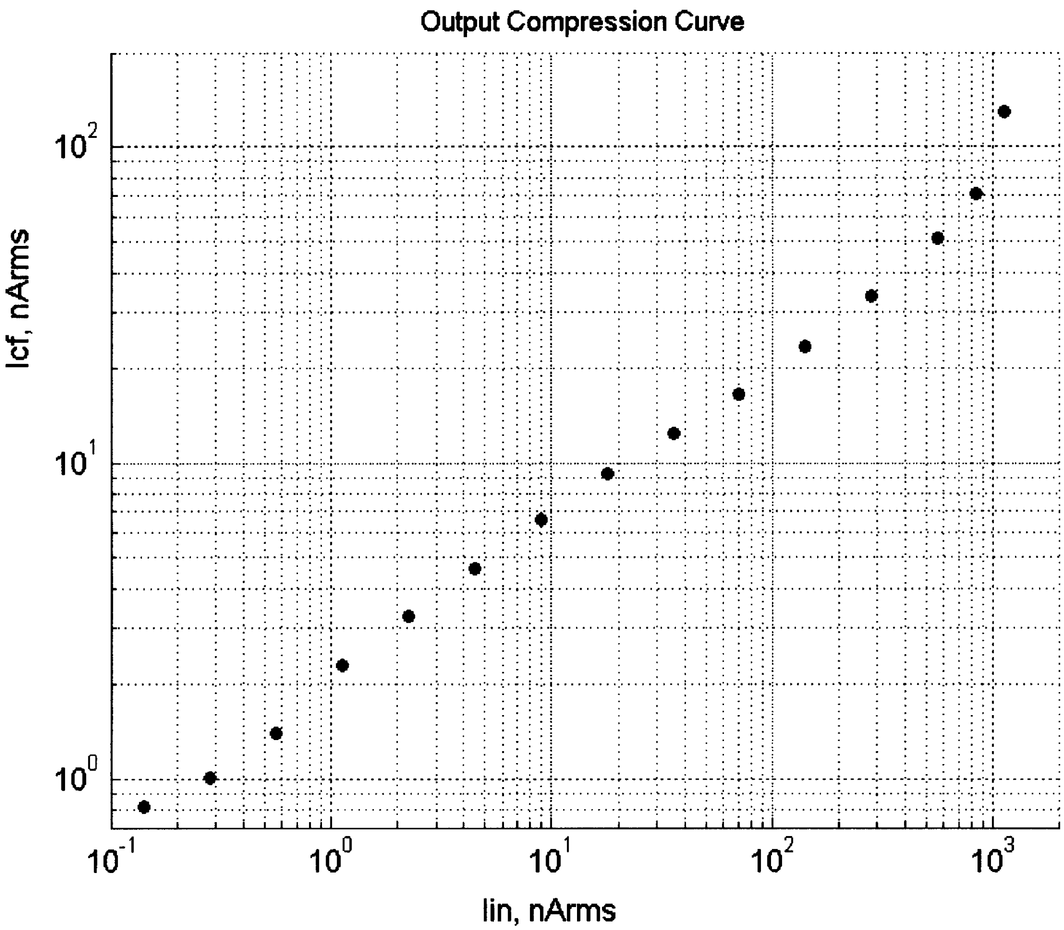


Figure 14: Compression curve taken at output 17. This figure shows the dependence of the output signal measured at the best frequency for the soft input signals versus the input signal level. The electronic cochlea compresses almost 4 decades of input signal magnitude span into approximately 2 decades of output signal variation, which is similar to the biological cochlea.

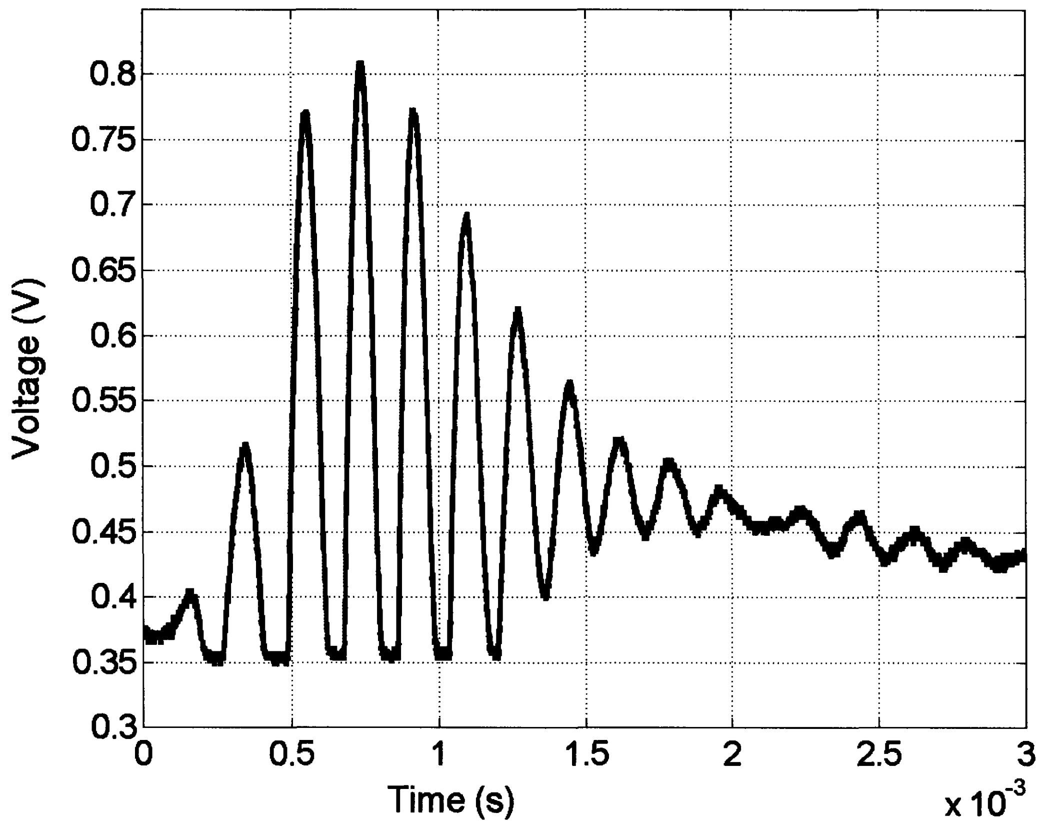


Figure 15: Step response taken from the half-wave rectified output 17. This figure shows that the step response peaks after only 4 periods of the local best frequency demonstrating low phase run-up and group delay accumulation in the cochlear cascade of all-pass filters.

6.4.2. The discrete version of the cochlear cascade with AGC

To correct the mistake in on-chip programming DACs, we have built a discrete version of the cochlear cascade with 30 stages.

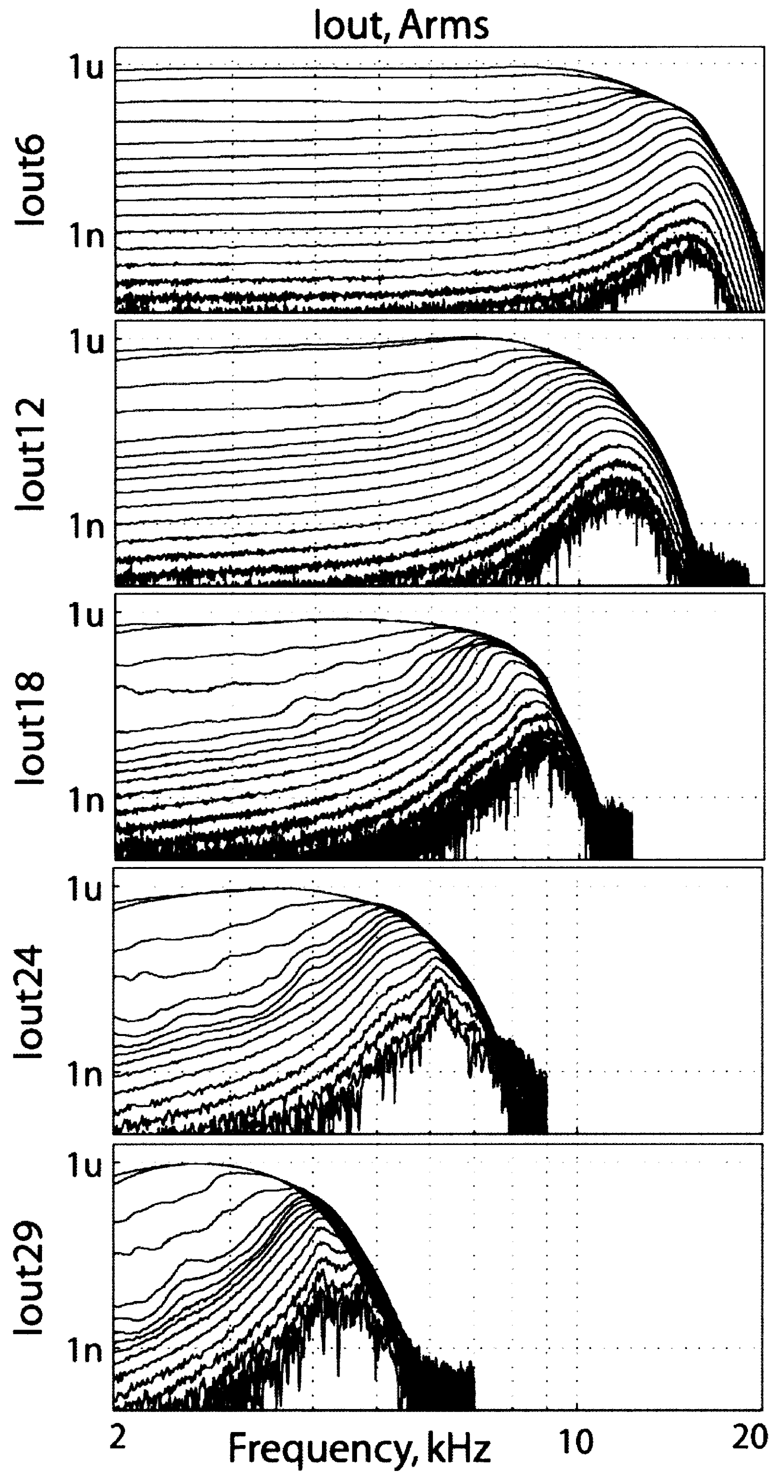


Figure 16: The discrete version of the cochlear cascade with AGC. Magnitudes of frequency responses taken at outputs 6, 12, 18, 24 and 29 with varying intensity of the input signal are shown. This figure demonstrates the spatial distribution of the peak frequencies. The peak gain at low intensities is approximately 50-dB, and Q_{-10dB} is approximately 4.2. The usable frequency range is 4-16 kHz.

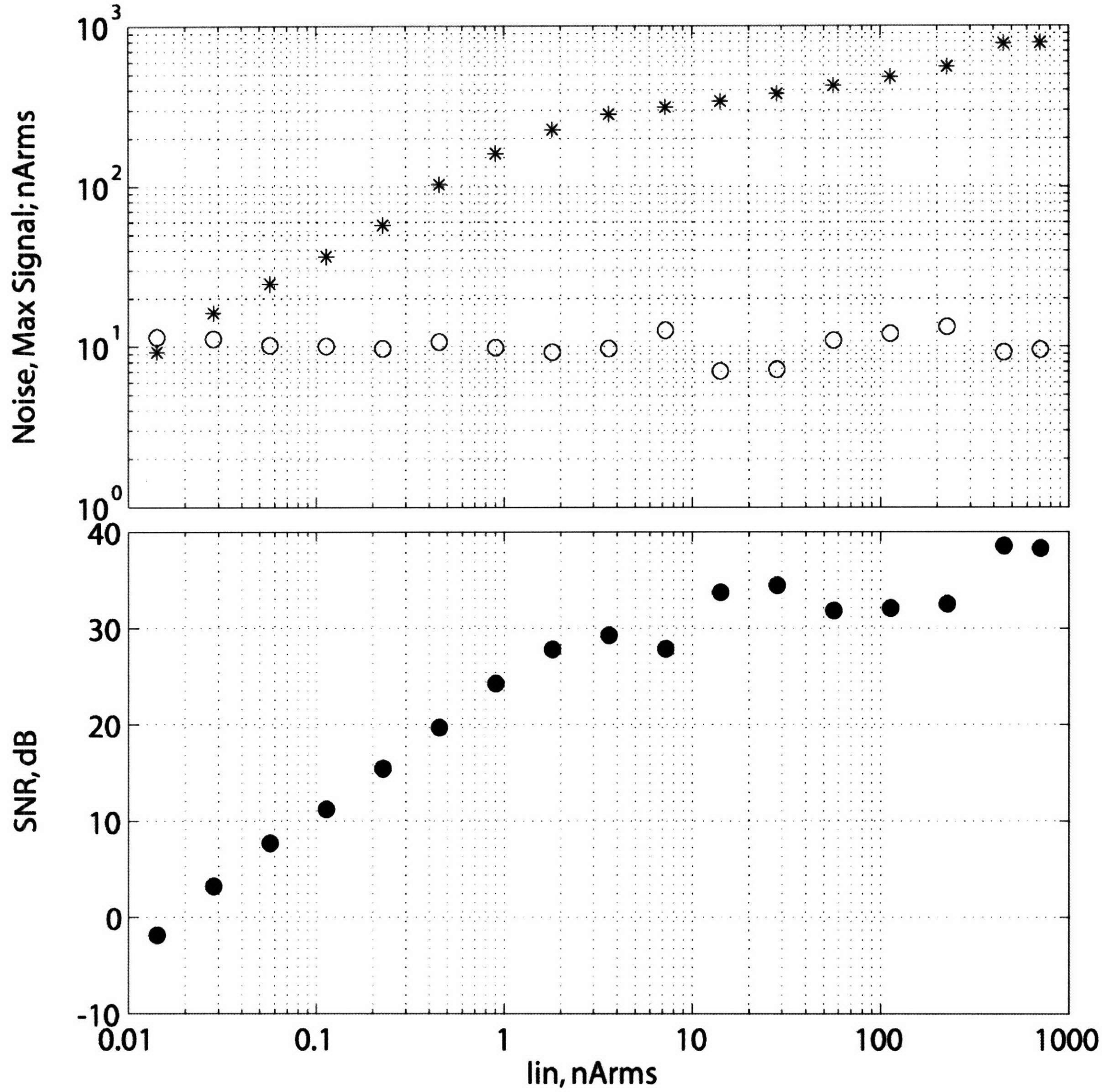


Figure 17: The discrete version of the cochlear cascade with AGC. Top figure shows the noise (o) and the maximum signal (*) and the bottom figure shows SNR calculated from the top figure as a function of the input signal level. Output 24 is shown. Similarly to the integrated version, two conflicting effects take place: the increase in noise due to increasing bias currents; and the decrease in the cochlear gain due to the gain control (compression) and the associated decrease in noise. This figure demonstrates that the dynamic range of the integrated version can be improved to 92-dB by redesigning the programming DACs on the chip to increase the stage gain to 6-dB and the cochlear cascade gain from 24-dB to 50-dB.

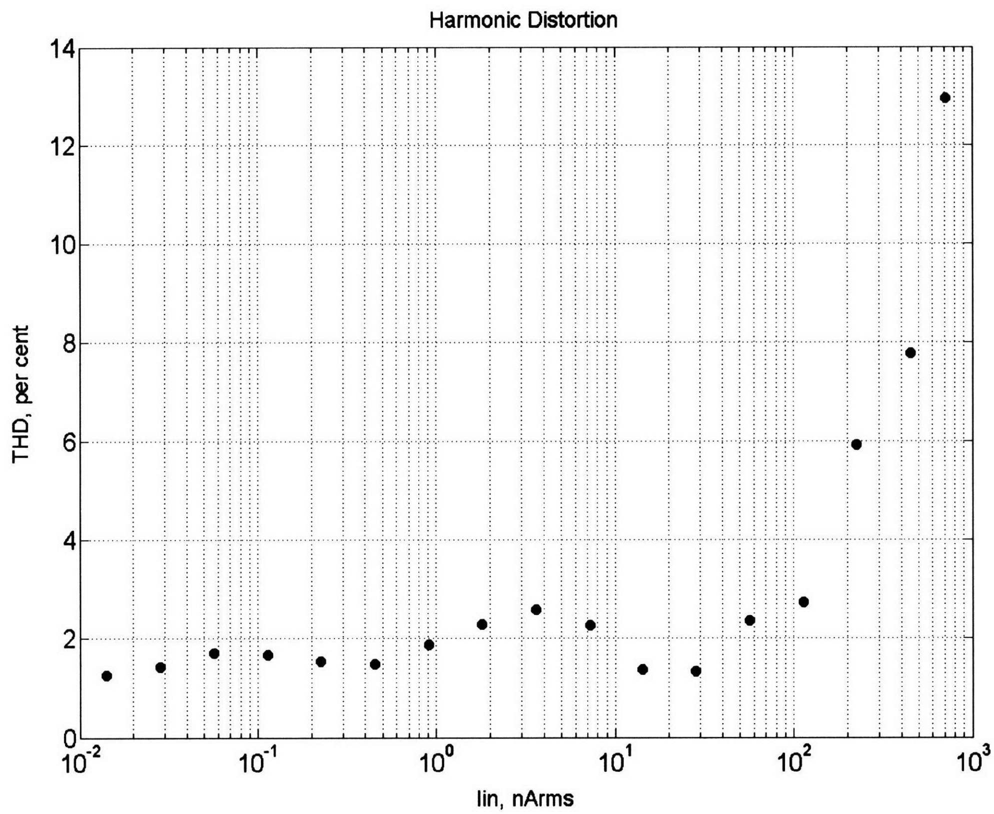


Figure 18: Harmonic distortion versus input signal level for output 24 of the discrete version of the cochlear cascade with AGC.

6.5. Comparison to State-Of-The-Art

Table I shows the comparison of our two designs to the best state-of-the-art:

	LPF cascade [Sarpeshkar 1997]	Cochlear chip	Discrete version
# of sections per octave	12	12	12
Frequency range (# of octaves)	100Hz-18kHz (7.4 octaves)	3.5kHz-14kHz (2 octaves)	4kHz-16kHz (2 octaves)
Peak gain	35 dB	24 dB	50 dB
Phase run-up	~7 cycles	~2 cycles	~2.5 cycles
Q_{-10dB}	1.5	3.6	4.2
High-freq. rolloff	~74 dB/octave	~100 dB/octave	~150 dB/octave

Input DR	61 dB	79 dB	92 dB
SNR (prevailing number)	23 dB	40 dB (44 dB max)	40 dB (44 dB max)
Quiescent power cons.	500 μW	41 μW	102 μW
Vdd	5 V	1 V	1.1 V

6.6. Summary

We have presented two electronic cochlea designs

6.7. References

[1] M.W. Baker "Analog Front End," JSSC

7. Conclusions

- Simple rational approximation to partition impedances shown to capture the cochlea's essential features. It achieves maximum gain in a minimum number of stages.
- The novel cascade of all-pass stages reduces phase lag and group delay, sharpens high-frequency roll-off slopes.
- A novel log-domain technique demonstrates a reduction in power consumption and increase in SNR by a factor of Q, and an increase in dynamic range by a factor of Q^4
- A 33-stage 0.18μm silicon cochlea achieves 79dB of dynamic range with 41μW power consumption on a 1V power supply over a usable frequency range of 3.5kHz-14kHz
  - An 18dB improvement in dynamic range and a 12.5x reduction in power consumption over state-of-the-art silicon cochleas

8. Future Work

Redesign the programming DACs to match the cochlear gain and achieve 92dB dynamic range of the discrete version with no increase in power consumption.

Incorporate the cochlea chip into a speech processor for cochlear implants: