

A Bio-Inspired Cochlear Heterodyning Architecture for an RF Fovea

Soumyajit Mandal, *Member, IEEE*, and Rahul Sarpeshkar, *Senior Member, IEEE*

Abstract—We discuss the use of cochlear models for spectrum analysis at radio frequencies. We describe performance characteristics of such models, including noise, dynamic range, and frequency resolution. We show that the addition of phase information improves frequency estimation as compared to the use of amplitude information alone. In particular, the use of both amplitude and phase information in a novel nonlinear bio-inspired center-surround coincidence-detection stage simultaneously improves frequency estimation and implements a lowpass-to-bandpass transformation on cochlear outputs. In order to further improve frequency estimation we propose a novel wireless receiver architecture that is a broadband generalization of narrowband heterodyning systems commonly used in radio. We term this architecture *cochlear heterodyning*. It exploits the efficiency of cochlear spectrum analysis to perform parallel, multi-scale analysis of wideband signals and can be constructed with cochlea-like traveling-wave structures. When combined with our prior work on an RF cochlea, such architectures may be useful in cognitive radios for creating *RF foveas* that select narrowband components present within wideband, but spectrally sparse signals. The operation of RF foveas is analogous to how the eye foveates on narrow but interesting portions of an image. Analogies between spectrum analysis and the process of successive-subranging analog-to-digital conversion illustrate how successively finer frequency resolution is achieved in an RF fovea. Finally, we show that RF foveas can be used in feedback loops to perform interference cancellation.

Index Terms—Cochlear heterodyning, cognitive radio, frequency estimation, frequency resolution, phase analysis, radio frequency (RF) cochlea, RF fovea, software radio, spectrum analysis.

I. INTRODUCTION

WE have recently demonstrated a single-chip radio-frequency (RF) spectrum analyzer based upon a model of the biological cochlea, or inner ear, that we termed an *RF cochlea* [1]. The RF cochlea extends previous work on biologically-inspired audio-frequency spectrum analysis [2]–[5] into the GHz range. The RF cochlea chips shown in Fig. 1 perform spectrum analysis over a broad range of RF frequencies (600 MHz to 8 GHz in this implementation) with low-power, wide dynamic range, low latency, and efficient hardware usage as discussed in [1].

Manuscript received October 01, 2010; revised December 09, 2010; accepted February 01, 2011. Date of publication March 28, 2011; date of current version June 29, 2011. This paper was recommended by Associate Editor Y. Lian.

S. Mandal was with the Massachusetts Institute of Technology (MIT), Cambridge, MA 02139 USA. He is now with Schlumberger-Doll Research, Cambridge, MA 02139 USA.

R. Sarpeshkar is with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology (MIT), Cambridge, MA 02139 USA (e-mail: rahuls@mit.edu).

Digital Object Identifier 10.1109/TCSI.2011.2123710

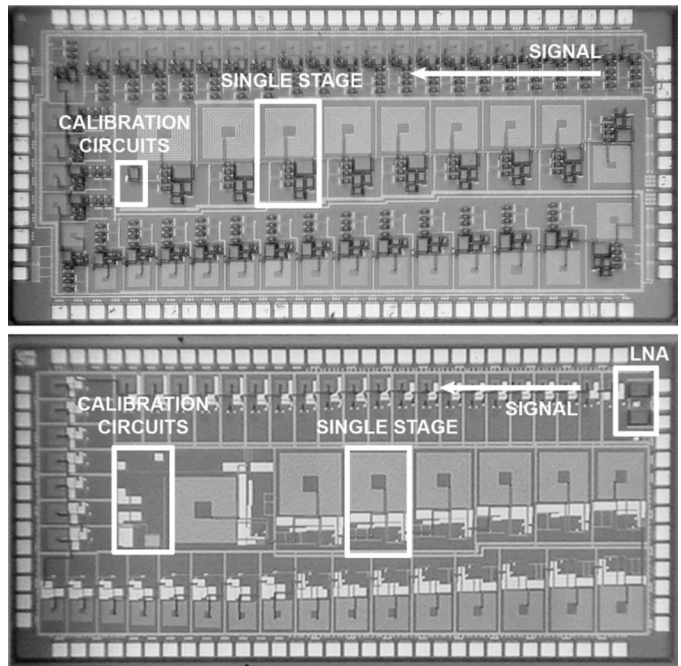


Fig. 1. (top) Die photographs of the bidirectional and (bottom) unidirectional RF cochlea chips described in [1]. Each chip was fabricated in the UMC 0.13- μm CMOS process and measures 3 mm \times 1.5 mm in size.

The RF cochlea provides real-time sensing of the RF spectrum, which is important for a range of wireless applications. The commercially interesting range of frequencies extends over approximately three orders of magnitude from 100 MHz to 100 GHz. Nevertheless, this natural resource is becoming increasingly crowded. At the same time, customers want higher quality of service, particularly increased reliability, and higher data rates. In order to satisfy these conflicting goals, radios must utilize the RF spectrum in a more efficient manner. The *cognitive radio* paradigm has been proposed as a solution to this important problem [6]–[8].

Cognitive radios add intelligence to traditional wireless communication systems. They are designed to sense their local RF environment, learn from it, and dynamically adapt their operating parameters, such as transmit power, carrier frequency and modulation strategy, appropriately. In particular, they actively search for spectrum holes, i.e., locations in the time-frequency plane that are suitable for communication, and operate there. As a result, they must be built upon dynamically reconfigurable hardware platforms, such as software-defined receivers and transmitters [9]–[13]. Such transceivers present numerous hardware design challenges that demand innovation at every step of the signal chain, including power amplifiers [14], antennas [15], mixers [16], filters [17]–[19], clocks [20],

spectrum sensors [21]–[23], and analog-to-digital converters [24].

Unused VHF and UHF TV channels (54–806 MHz in the United States) currently provide most of the spectrum holes that cognitive radios seek to exploit [25]. As a result, the recent IEEE 802.22 cognitive radio standard focuses on this part of the wireless spectrum. However, in the future a much wider range of frequencies will need to be monitored for potentially usable holes, making the RF cochlea attractive as a rapid, broadband spectrum sensor.

In this paper, we first analyze noise, dynamic range, and frequency-resolution limits of cochlea-like traveling-wave architectures such as the RF cochlea. This analysis shows that the use of phase information significantly improves frequency resolution. We suggest two phase-based lowpass-to-bandpass transformations on the RF cochlear outputs to improve the frequency resolution of the RF cochlea. We also develop a novel multi-resolution spectrum-analyzer architecture based on the RF cochlea, which effectively implements a broadband generalization of heterodyning. We refer to this architecture as a *cochlear heterodyning* architecture. When combined with our prior work on an RF cochlea, such architectures may be useful in cognitive radios for creating *RF foveas* that select the narrowband components present within wideband, but spectrally sparse signals. We show that analogies between spectrum analysis and the process of successive-subranging analog-to-digital conversion illustrate how successively finer frequency resolution is achieved in an RF fovea. We also show that RF foveas can be used in feedback loops to perform interference cancellation. Our architectures can benefit software-defined receivers by providing real-time estimates of interesting portions of the local RF spectrum with scalably high precision and low hardware and power costs.

This paper is organized as follows. In Section II, we analyze the performance of cochlea-like traveling-wave architectures such as the RF cochlea. In Section III, we describe lowpass-to-bandpass transformation techniques on cochlear outputs that are useful for spectrum analysis. In Section IV, we first describe how the problem of frequency estimation is analogous in many ways to analog-to-digital conversion. We then describe the concepts behind cochlear heterodyning and show how it can be used to build an efficient multi-resolution spectrum analyzer that we refer to as the *RF fovea*. In Section V, we conclude the paper by summarizing our contributions. A brief version of our work has been presented at a conference [26]. In this invited journal submission, we expand on and enhance the material in the conference version.

II. ANALYSIS OF THE RF COCHLEA

Cochlear models form the basis of our frequency estimation algorithms. To first order, the cochlea can be modeled as a 1-D transmission line with properties that scale exponentially with position. A particular position on the line responds selectively to a characteristic, or best frequency that decreases exponentially with distance from the input terminal. This property generates a frequency-to-place mapping that forms the basis of cochlear spectrum analysis.

Cochlear models can be implemented in hardware at radio frequencies by assuming linearity and using the Wentzel-Kramers-Brillouin (WKB) approximation to solve the resultant spatially-varying wave equation on the transmission line. This step is followed by spatial discretization and appropriate scaling of element values. Spatial discretization refers to the process by which the continuous transmission line is approximated by a discrete line with N_{nat} lumped stages per e-fold in frequency range, or $N_{\text{oct}} = N_{\text{nat}} \ln(2)$ stages per octave. We refer to this architecture as the *bidirectional cochlea*. The propagation of backward, or reflected, waves in the cochlea can usually be neglected. In this case small sections of the transmission line can be approximated as unidirectional filters, resulting in an alternate architecture known as the *unidirectional cochlea*. It consists of a cascade of filters that are frequency-scaled versions of a single normalized prototype. We have implemented both models in integrated-circuit form at RF [1], as shown in Fig. 1.

In this section, we describe important properties of the unidirectional cochlea. With relatively minor differences, which are summarized in the last subsection, our conclusions are applicable to the bidirectional cochlea as well [27].

A. Transfer Functions

The transfer function of the unidirectional cochlea after the n th stage is the product of the transfer functions of the first n cochlear filters, i.e., $TF(\omega, n) = \prod_{i=1}^n H_i(\omega, n)$. The values of TF at a fixed frequency ω form a set that is known as the *spatial response* of the cochlea.

The transfer functions of exponentially-scaled transmission lines or filter cascades, such as our cochlear models, can be expressed in terms of a single normalized frequency variable ω_n , instead of frequency ω and position n separately [28]. Here $\omega_n = \omega/\omega_c(n) = \exp(n/N_{\text{nat}})(\omega/\omega_c(0))$, where $\omega_c(n) = \omega_c(0) \exp(-n/N_{\text{nat}})$ is the characteristic frequency of the n th stage. This is an interesting and important property. In log-frequency space the normalized frequency is given by

$$\ln(\omega_n) = \frac{n}{N_{\text{nat}}} + \ln\left(\frac{\omega}{\omega_c(0)}\right). \quad (1)$$

The space and frequency-dependent terms add linearly in this combined variable. Thus cochlear spatial responses (constant ω), when normalized by N_{nat} , are identical to cochlear frequency responses (constant n) when plotted on a logarithmic frequency scale. Similarly, spatial and frequency derivatives of cochlear transfer functions are proportional to each other

$$\frac{\partial(TF)}{\partial \ln(\omega)} = N_{\text{nat}} \frac{\partial(TF)}{\partial n}. \quad (2)$$

The normalized transfer function of each filter in our unidirectional cochlea was derived by approximating the WKB solution [1]. It is given by

$$H_n(s_n) = \frac{s_n^2 + 2ds_n + 1}{(1 + \sqrt{Q_{\text{line}}/\mu})s_n^2 + (2d + \sqrt{Q_{\text{line}}})s_n + 1} \quad (3)$$

where $s_n = j\omega_n$ is the normalized frequency variable, and d , μ , and Q_{line} , along with N_{nat} , are design parameters. The function $|H(s_n)|$ has a low-frequency value of 1, a maximum value of

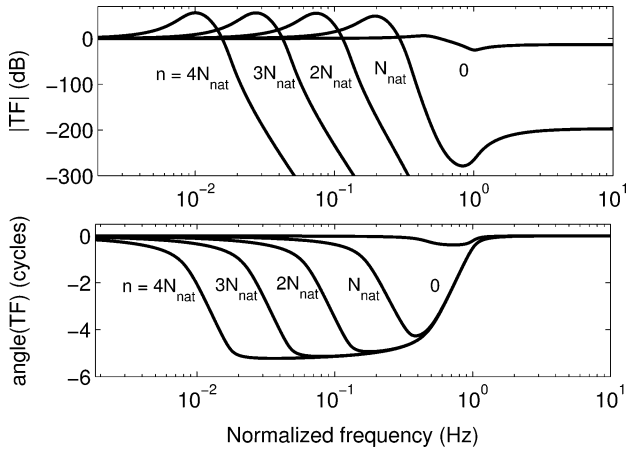


Fig. 2. Typical transfer functions of the unidirectional cochlea for stages $n = N_{\text{nat}}$ apart. Phase is shown in cycles, i.e., units of 2π radians. Parameters used to draw this plot were: $N_{\text{nat}} = 14$, $d = 0.1$, $\mu = 0.2$, and $Q_{\text{line}} = 0.5$.

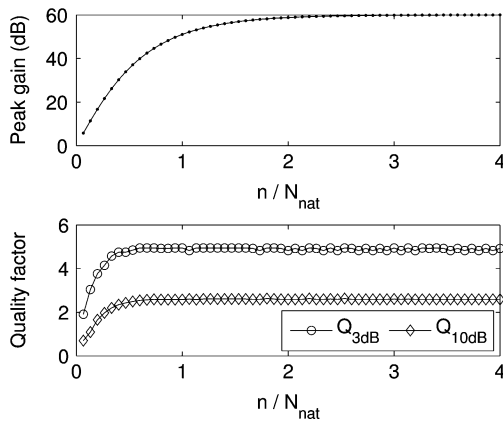


Fig. 3. Behavior of the unidirectional cochlea transfer functions as a function of position n along the cascade. Top: peak gain; bottom: 3 and 10 dB quality factors. Parameters used to draw this plot were: $N_{\text{nat}} = 14$, $d = 0.1$, $\mu = 0.2$, and $Q_{\text{line}} = 0.5$.

$G_{\text{stage}} > 1$, and asymptotes to $1/(1 + \sqrt{Q_{\text{line}}}/\mu) < 1$ at high frequencies. This filter function has several advantages, such as reduced group delay and improved frequency resolution, over other cochlear filters described in the literature [1]. However, the results that follow are qualitatively independent of the precise choice of cochlear filter.

Cochlear transfer functions are characterized by peak gain $G(n)$ and quality factors $Q_{3\text{ dB}}(n)$ and $Q_{10\text{ dB}}(n)$ (defined as the center frequency divided by the 3 or 10 dB bandwidth, respectively). Because of exponential scaling of center frequencies, each cochlear transfer function may be approximated as the product of N_{nat} identical filter transfer functions. Thus we expect $G(n)$ for $n > N_{\text{nat}}$ to be constant and given by $G \approx (G_{\text{stage}})^{N_{\text{nat}}}$. Since $G_{\text{stage}} > 1$ the peak gain increases rapidly with N_{nat} . This behavior is independent of the precise shape of the filter transfer function. Typical unidirectional cochlear transfer functions are shown in Fig. 2 for different values of n . Bidirectional models have qualitatively similar behavior.

Fig. 3 shows simulated values of $G(n)$, $Q_{3\text{ dB}}(n)$ and $Q_{10\text{ dB}}(n)$ for $N_{\text{nat}} = 14$, and typical filter parameters ($d = 0.1$, $\mu = 0.2$, and $Q_{\text{line}} = 0.5$). The peak filter gain

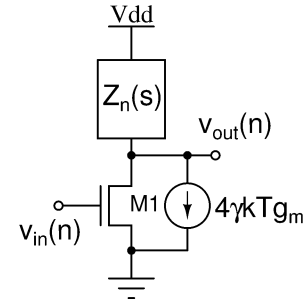


Fig. 4. Simplified cochlear filter circuit, used for noise calculations.

$G_{\text{stage}} \approx 1.95$ (5.8 dB) for these parameter values. We see that $G(n)$ asymptotes to a fixed value for $n > 2N_{\text{nat}}$. Similarly, the quality factors asymptote for $n > N_{\text{nat}}/2$. This asymptoting behavior occurs because of exponential scaling, and is insensitive to the values of d , μ , and Q_{line} .

We define the asymptotic gain G and quality factor $Q_{3\text{ dB}}$ of the cochlear transfer functions as their values at large n . Simulations show that $G \approx (0.8G_{\text{stage}})^{N_{\text{nat}}}$, while $Q_{3\text{ dB}}$ increases approximately as $\sqrt{N_{\text{nat}}}$. These quantities also vary as a function of d , μ , and Q_{line} for a fixed value of N_{nat} . Decreasing d and μ cause G_{stage} to increase, while Q_{line} has much less effect on G_{stage} . Thus, both G and $Q_{3\text{ dB}}$ increase sharply as d and μ decrease, but are approximately independent of Q_{line} .

B. Noise

Each cochlear filter is assumed to be strictly unidirectional. The n th stage filters the output noise voltage $v_{\text{ns}}(n-1)$ produced by the previous, $(n-1)$ th stage, and adds some noise of its own, so we get

$$\overline{v_{\text{ns}}^2(n)} = \overline{v_{\text{ns}}^2(n-1)} |H_n(j\omega)|^2 + \overline{v_n^2} \quad (4)$$

where v_n is the noise PSD produced by the n th stage, and $H_n(j\omega)$ is its transfer function. Because H_n is an exponentially-frequency scaled version of a common prototype H_1 , we have $H_n(\omega) = H_1(\omega e^{-(n-1)/N_{\text{nat}}})$. For simplicity, we assume that all the filtering within each stage is provided by a passive impedance $Z_n(s)$ which contributes a negligible amount of noise to v_n . A single transistor M_1 , shown in Fig. 4, provides buffering and voltage gain. It converts input voltage into current such that it can be filtered by $Z_n(s)$, and is modeled as an ideal transconductance of value g_m . The value of g_m is assumed to be the same for each filter.

There is only one active component in the filter circuit shown in Fig. 4, namely the transistor M_1 . The linear range of M_1 , given by $V_L = I/g_m$, is approximately equal to that of the filter. Here I is the DC bias current flowing through M_1 . The value of V_L is independent of I when M_1 operates in subthreshold, and is given by ϕ_T/κ [4]. It increases proportionately with \sqrt{I} above threshold.

Ignoring $1/f$ noise, each transistor adds a white noise current with PSD = $4\gamma kTg_m$ to the output node, where γ , known as the excess noise factor, has a value of $2/3$ for long-channel transistors. Practical on-chip cochlear filters will be somewhat different from the circuit shown in Fig. 4 [1]. However, we shall study this simplified circuit, because its noise properties are very

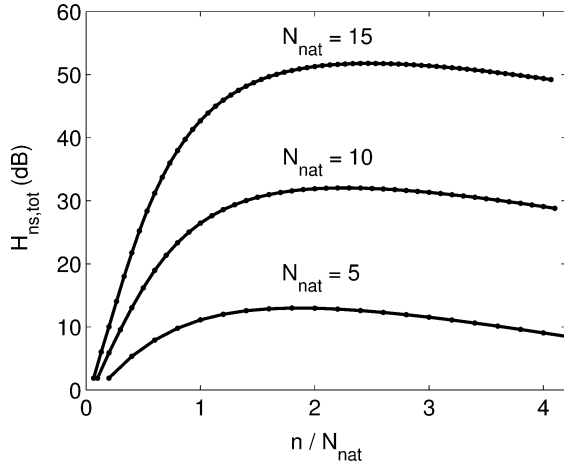


Fig. 5. Effective noise magnitude transfer function in the unidirectional cochlea as a function of position n along the cascade for various values of N_{nat} . Parameters fixed for this plot were: $d = 0.1$, $\mu = 0.2$, and $Q_{\text{line}} = 0.5$.

similar to more practical circuits, while requiring considerably less algebra to analyze.

The voltage PSD contributed by the simplified filter circuit to its output node is given by

$$\overline{v_n^2} = 4\gamma kT g_m |Z_n(j\omega)|^2 = \frac{4\gamma kT}{g_m} |H_n(j\omega)|^2 \quad (5)$$

where $H_n(s) = g_m Z_n(s)$ is the filter's transfer function. Combining (4) and (5), we get the net PSD at the output of the n th filter

$$\overline{v_{\text{ns}}^2(n)} = \left[\overline{v_{\text{ns}}^2(n-1)} + \frac{4\gamma kT}{g_m} \right] |H_n(j\omega)|^2. \quad (6)$$

We can use the fact that $v_{\text{ns}}(0) = 0$ (i.e., the input is noiseless) and the iteration formula shown in (6) to explicitly derive $\overline{v_{\text{ns}}^2(n)}$

$$\overline{v_{\text{ns}}^2(n)} = \frac{4\gamma kT}{g_m} \times \sum_{i=1}^n \prod_{j=n+1-i}^n |H_j(j\omega)|^2 \equiv \frac{4\gamma kT}{g_m} H_{\text{ns},n}(j\omega) \quad (7)$$

where $H_{\text{ns},n}(j\omega)$ is defined as the noise transfer function of the n th stage. These functions are similar, but not identical, to the squared magnitude of the signal transfer functions. The total noise voltage can be found by integrating the PSD over all frequencies

$$\overline{v_{\text{tot}}^2(n)} = \frac{4\gamma kT}{g_m} \int_0^\infty H_{\text{ns},n}(f) df \equiv \frac{4\gamma kT}{g_m} H_{\text{ns,tot}} \quad (8)$$

where $2\pi f = \omega$. The quantity $H_{\text{ns,tot}}$, which is also a function of n , has dimensions of bandwidth, and is plotted for typical filter parameters and various values of N_{nat} in Fig. 5.

The variation in total noise voltage with position n is largely independent of d , μ , and Q_{line} . The noise PSD increases with increasing n because all stages before it contribute noise, but the noise bandwidth decreases as $\exp(-n/N_{\text{nat}})$. As a result the total noise voltage, which is proportional to $\sqrt{H_{\text{ns,tot}}}$, increases with n , reaches a maximum around $n = 2N_{\text{nat}}$, and then slowly decreases. However, the maximum is very gentle such that like $G(n)$ the noise may also be approximated as being nearly constant with n for $n > 2N_{\text{nat}}$, as found in prior audio-frequency

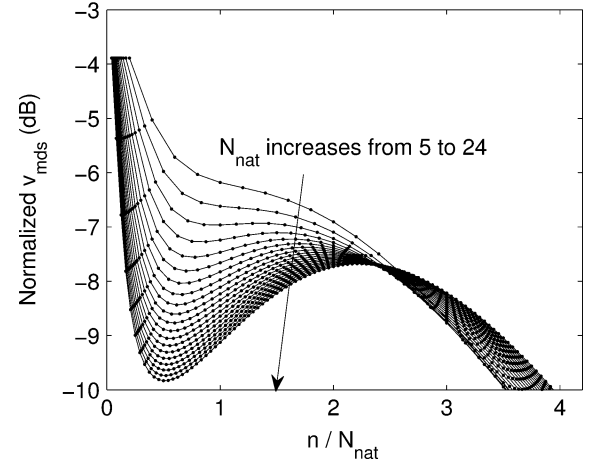


Fig. 6. Normalized minimum detectable signal in the unidirectional cochlea as a function of position n along the cascade for various values of N_{nat} . Parameters used to draw this plot were: $d = 0.1$, $\mu = 0.2$, and $Q_{\text{line}} = 0.5$.

cochlea designs [2]. Because this behavior is determined by exponential scaling, it is largely independent of the particular choice of filter parameters.

The *minimum detectable signal*, $v_{\text{mds}}(n)$ is defined as the amplitude of the smallest input signal that just exceeds the output noise of the n th stage. The value of $v_{\text{mds}}(n)$ reaches a minimum when the frequency of this signal is equal to the center frequency of the n th stage. In this case, we have $v_{\text{mds}}(n) = v_{\text{tot}}(n)/G(n)$.

Since $v_{\text{tot}}(n) \propto 1/\sqrt{g_m}$, we can burn power to increase g_m , lower the minimum detectable signal and increase sensitivity. For the purpose of comparing cochlear designs with different parameter values we can eliminate this dependence on g_m by defining a normalized version of $v_{\text{mds}} = \sqrt{H_{\text{ns,tot}}}/G(n)$. Fig. 6 plots this quantity as a function of n for various values of N_{nat} . Changing N_{nat} leaves the minimum detectable signal unchanged to within ~ 4 dB, because N_{nat} affects the signal and noise transfer functions in similar ways, but leaves their ratio, which sets v_{mds} , fixed. However, when examined in finer detail, v_{mds} does vary in a complex way with n , in general decreasing slowly as n increases. This behavior occurs because the signal and noise transfer functions vary in similar, but not identical, ways with spatial position n .

C. SNR and Dynamic Range

The range of signals that can be handled by the unidirectional cochlea is limited at the low end by noise, and at the high end by the linear range of the filters within it. All cochlear models use gain control to increase dynamic range. In the presence of gain control, the gain seen by a small-amplitude tone (frequency component) will be reduced by the simultaneous presence of a large-amplitude tone, because the latter will dominate the total amplitude, which usually determines the amount of gain compression. Thus, the sensitivity of the system to the smaller tone will be diminished. This phenomenon is known as *tone-to-tone suppression*, and, in the auditory system, causes a psycho-acoustic phenomenon known as *masking*. The traveling wave structure of the cochlea leads to *asymmetric masking*: Low-frequency tones pass through most of the cochlea before peaking, and therefore suppress the gain of the filters that are

tuned to higher frequencies. On the other hand, high-frequency tones peak early and are then filtered out, and therefore cannot affect the gain of later filters tuned to lower frequencies. As a result of masking, frequency discrimination worsens when the two signals have very different amplitudes.

The maximum undistorted signal amplitude at the output of the n -filter is V_L . As a result, the maximum signal-to-noise ratio (SNR) of the cochlear outputs is given by

$$\text{SNR}_{\max}(n) = \left(\frac{V_L}{v_{\text{tot}}(n)} \right)^2. \quad (9)$$

As long as the cochlea remains completely linear the dynamic range (DR) of input signals it can handle is equal to SNR_{\max} . However, gain control can be used to increase input-referred dynamic range [2], [3]. Gain reduction at high input amplitudes reduces harmonic distortion at the cost of reduced sensitivity to small signals.

The peak gain $G(n)$ of the cochlear transfer functions can be reduced for large input signals. This strategy is a form of distributed, or parallel gain control that uses the filtering provided by the cochlea to compress different frequency components by different amounts. An important advantage of the filter transfer function shown in (3) is that it allows gain control to be implemented in a very simple way: the single parameter d can be increased with local signal amplitude $|A|$, thus decreasing G_{stage} for large signals. For example, we can increase d linearly with A , i.e., as $d = d_{\min} + \sigma|A|$, where σ is a constant. It can be shown, as long as $d \ll 1$, that this purely local technique keeps the timing of zero-crossings in the cochlear impulse responses invariant with $|A|$ [29].

We can define the dynamic range of the cochlea by supposing that, at the maximum acceptable level of gain compression the gain of the n th stage is reduced from $G(n)$, its value for small signals, to a new value $G_{\text{comp}}(n)$. The maximum allowable input signal amplitude is then increased from $V_L/G(n)$ to $V_L/G_{\text{comp}}(n)$, while the minimum detectable signal is unchanged, and given by $v_{\text{tot}}(n)/G(n)$. Thus the input-referred dynamic range becomes

$$\text{DR} = \left(\frac{V_L/G_{\text{comp}}(n)}{v_{\text{tot}}(n)/G(n)} \right)^2 = \text{SNR}_{\max} \left(\frac{G(n)}{G_{\text{comp}}(n)} \right)^2. \quad (10)$$

Equation (10) is only valid for the largest tone in the input spectrum. The simultaneous presence of other large tones will reduce $G(n)$, and hence the dynamic range.

In RF systems the limit of linear operation is often defined as the 1 dB compression point, i.e., the input amplitude when $G_{\text{comp,dB}} = G_{\text{dB}} - 1$ dB. This is an overly pessimistic limit for the cochlea, which is fundamentally nonlinear. A convenient upper bound for DR, which we will henceforth denote as DR, is obtained by setting $G_{\text{comp}}(n) = 1$, i.e., assuming that there is no peaking in the cochlear transfer functions. Fig. 7 shows simulated values of SNR_{\max} and DR as a function of position n for various values of N_{nat} and typical values of other parameters. We see that SNR_{\max} decreases as N_{nat} decreases, because the noise transfer functions increase while V_L remains fixed. However, DR is largely independent of N_{nat} , because $\text{DR} = (V_L/v_{\text{mds}}(n))^2$, and, as shown by Fig. 6, $v_{\text{mds}}(n)$ is

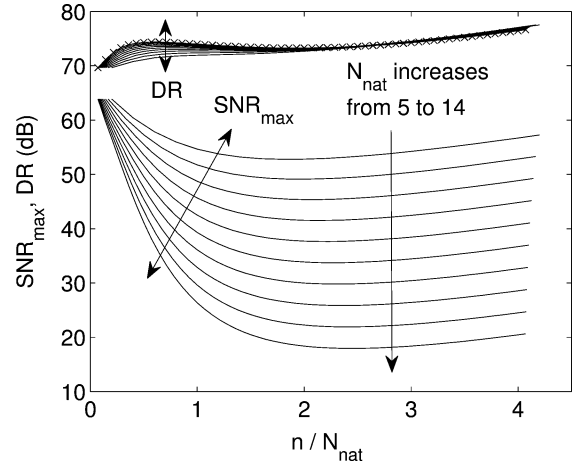


Fig. 7. Maximum SNR and input-referred DR of the unidirectional cochlea as a function of position n along the cascade for various values of N_{nat} . Parameters fixed for this plot were: $d = 0.1$, $\mu = 0.2$, $Q_{\text{line}} = 0.5$, $\gamma = 1$, $g_m = 1/60\Omega$, $V_L = 150$ mV and a maximum cochlea frequency of 6 GHz.

approximately independent of N_{nat} . For the typical parameter values used to draw Fig. 7, the average value of v_{mds} is approximately 50 μV .

Fig. 7 shows that we have to rely on increasing amounts of gain compression to get the full dynamic range available from the cochlea as N_{nat} increases. We expect similar behavior if the filter parameters d and μ are decreased for a fixed value of N_{nat} , because in either case G_{stage} increases, causing more peak gain in both the signal and noise transfer functions and reducing SNR_{\max} .

D. Frequency Resolution

SNR_{\max} is important because it limits the maximum frequency resolution of the cochlea. Consider a single input tone $S(A, \omega)$ with a amplitude of A and a frequency ω . We can measure A by using an envelope detector (ED) at the input of the cochlea. In order to find ω we measure the amplitude and/or phase of the cochlear outputs. The output of the n th stage, S_n is given by

$$S_n(A, \omega) = S(A, \omega)TF_n(\omega) = S(A, \omega)G_n(\omega)e^{j\phi_n(\omega)} \quad (11)$$

where $G_n(\omega)$ and $\phi_n(\omega)$ are, respectively, the amplitude and phase of the cochlear transfer function TF_n at the input frequency ω . From (11), the sensitivity of the cochlear output to ω is given by

$$\frac{dS_n(\omega)}{d\omega} = S(A, \omega) \left[\frac{dG_n}{d\omega} e^{j\phi_n} + jG_n e^{j\phi_n} \frac{d\phi_n}{d\omega} \right]. \quad (12)$$

Since $G_n(\omega)$ and $\phi_n(\omega)$ are known quantities, we can use amplitude and/or phase sensitive detectors at the cochlear outputs to estimate ω . Any measurement of S_n is corrupted by noise, making our estimate of ω uncertain and limiting the cochlea's frequency resolution. Assuming that the output noise has a magnitude of σ_s (RMS), the frequency resolution is

$$\sigma_\omega \approx \frac{\sigma_s}{\frac{dS_n}{d\omega}} = \frac{1}{\sqrt{2}\sqrt{\text{SNR}} \left[\frac{d \ln(G_n)}{d\omega} + j \frac{d\phi_n}{d\omega} \right]} \quad (13)$$

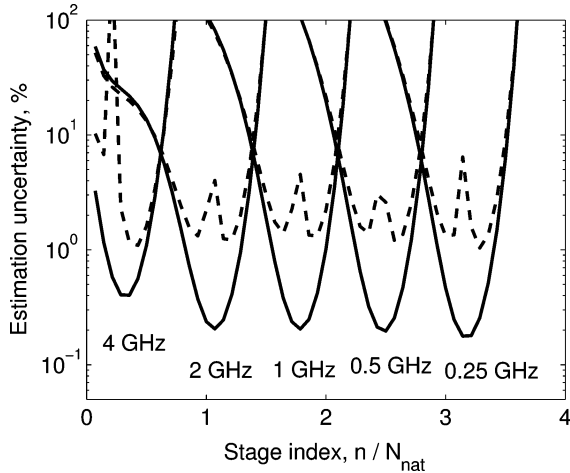


Fig. 8. Calculated frequency estimation error of the RF cochlea for amplitude and phase-based detectors (dashed and solid lines, respectively). The input amplitude was $500 \mu\text{V}$, corresponding to a maximum SNR of ~ 20 dB. Other parameters used to draw this plot were: $f_c(0) = 6$ GHz, $N_{\text{nat}} = 14$, $d = 0.1$, $\mu = 0.2$, and $Q_{\text{line}} = 0.5$.

where $\text{SNR} = |S_n|^2 / (2\sigma_s^2)$ is the output signal-to-noise ratio, with the factor of 2 being due to the usual difference between peak and RMS energies of a sinusoid. For an amplitude-sensitive detector, such as an ED, the second term in the denominator of (13) will be zero. Similarly, for a phase-sensitive detector such as a saturated mixer or firing-rate-saturated phase-locked neuron [30] the first term will be zero. The fractional frequency resolution σ_ω / ω in either case is shown in Fig. 8 for several input frequencies. Typical cochlear parameter values were assumed.

Fig. 8 shows that phase-based detection outperforms amplitude-based detection. Intuitively, SNR is highest near the peaks of the cochlear transfer functions. In this region G_n changes slowly with frequency because we are near a maximum (see Fig. 2). As a result, the sensitivity of amplitude-based detectors is relatively low. On the other hand, ϕ_n changes rapidly in this region, thus increasing the sensitivity of phase-based detectors. It has been shown that timing-synchrony features in cochlea-like filter bank outputs are more robust in the presence of noise than amplitude features [31]. In addition, Fig. 8 shows that the minimum estimation uncertainty decreases as the input frequency decreases from 4 to 2 GHz, and then remains constant. This behavior occurs because $|d \ln(G_n) / d\omega|$ and $|d\phi_n / d\omega|$ initially increase with n before reaching asymptotic values for $n > N_{\text{nat}}$, as seen in Fig. 2.

We have assumed that a single cochlear output is being used to estimate the input frequency. However, the cochlea provides many, highly overlapping outputs, and the information from several outputs can be combined to reduce estimation error. In fact, overlapping filter banks have been shown to be a computationally-efficient way to perform frequency estimation [32]. In the simplest case the estimates from multiple outputs can simply be averaged. However, the optimal strategy is to weigh the outputs by $1/\sigma_\omega$, i.e., the expected precision of the measurement, before averaging. This strategy is analogous to a matched filter. The expected values of σ_ω can be determined by utilizing any prior information that may be available about the input signal.

When the frequencies of the input signals are well-separated, a very simple fitting procedure known as *peak picking* can be used for coarse frequency estimation. In this technique, each amplitude peak in the cochlear spatial response is assumed to correspond to a different signal. The known frequency-to-space mapping of the cochlea is used to transform the measured spatial locations of the peaks into frequencies. The center frequencies of adjacent stages are separated by $\sim f_c / N_{\text{nat}}$, where f_c is the center frequency. Thus peak picking can localize the frequency corresponding to a spatial-response peak to a worst-case fractional bandwidth of $\Delta f_c / f_c \approx 1 / N_{\text{nat}}$. If the unknown frequency is uniformly distributed in this interval, the resulting RMS estimation error is

$$\sigma_0 = \frac{1}{\sqrt{12}} \left(\frac{\Delta f_c}{f_c} \right) \approx \frac{1}{\sqrt{12} N_{\text{nat}}} \quad (14)$$

E. Bidirectional RF Cochlea

In this section we briefly describe some unique features of the bidirectional cochlea. Our design consists of a lumped transmission line with exponentially-decreasing characteristic frequency. The normalized series impedance Z_n and shunt admittance Y_n of each section determine cochlear properties. Zweig derived these functions by fitting measurements of biological cochleas [28], but they were non-rational and hence difficult to implement. We found simple rational approximations of Zweig's functions and used them in our design [1]. They are given by

$$Z_n(s_n) = \frac{Q_{\text{line}} s_n \left(s_n^2 + \frac{\mu}{Q} s_n + \mu^2 \right)}{\mu^2 (s_n^2 + 2d s_n + 1)^2}, \quad Y_n(s_n) = s_n \quad (15)$$

where μ , Q , d , and Q_{line} are parameters. Intuitively, d controls the peak magnitude of Y_n , while μ controls the group delay. Assuming linearity, the spatial responses of the cochlea can be calculated in sinusoidal steady state by using the finite-difference equations for voltage and current on a lumped transmission line

$$\begin{aligned} I(n+1) &= I(n) - V(n)Y(n) \\ V(n+1) &= V(n) - I(n+1)Z(n) \\ V_{\text{out}}(n) &= V(n+1) - V(n) \end{aligned} \quad (16)$$

where $V(n)$, $I(n)$ are the line voltage and current at the n th stage, $V_{\text{out}}(n)$ is the differential voltage across the n th stage, and $Z(n)$ and $Y(n)$ are the series impedance and shunt admittance within this stage. This set of equations can be solved iteratively once the ratio $V(0)/I(0)$ is set by a known impedance.

The impedance at any position n , denoted by $Z_{\text{in}}(n)$, can also be found iteratively. We decompose the transmission line into two segments, one extending from location n to the apex (low-frequency end), and the other from n to the base (input, or high-frequency end). The input impedances of these segments are given by $Z_{\text{in}+}(n)$ and $Z_{\text{in}-}(n)$, respectively, and can be found using the following equations:

$$\begin{aligned} Z_{\text{in}+}(n+1) &= Z(n) + \left(Y(n) + \frac{1}{Z_{\text{in}+}(n)} \right)^{-1} \\ Z_{\text{in}-}(n-1) &= \left(Y(n-1) + \frac{1}{Z(n-1) + Z_{\text{in}-}(n)} \right)^{-1} \end{aligned} \quad (17)$$

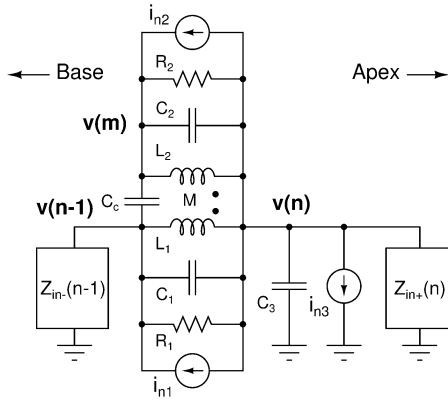


Fig. 9. Simplified equivalent circuit of a bidirectional cochlea stage used for noise calculations.

The equations above can be solved when the termination impedances at the base and apex of the cochlea, i.e., $Z_{in-}(1)$ and $Z_{in+}(N)$, are known, where N is the total number of stages. The impedance at n is $Z_{in}(n) = Z_{in-}(n) || Z_{in+}(n)$. The input impedance of the entire cochlea, i.e., $Z_{in+}(1)$, is denoted by Z_{in} for convenience and usually designed to be 50Ω [1].

The transfer functions of the bidirectional cochlea are largely similar to those of the unidirectional cochlea. Differences arise from the bidirectional nature of the structure, in particular because of reflected waves that propagate backwards. In general, the bidirectional cochlea transfer functions have lower peak gains, but higher quality factors. In addition, the peak gains are invariant with N_{nat} , while the quality factors increase slightly. The high frequency roll-off slope of the transfer functions is $20N_{nat}$ dB/decade, i.e., equal to that of a filter of order N_{nat} . At very low values of N_{nat} they are not smooth functions of frequency, but exhibit ripple because of inter-stage reflections.

In the bidirectional cochlea both peak gain and quality factor increase with spatial position n before saturating to constant values for $n > 0.3N_{nat}$, i.e., much faster than in the unidirectional cochlea. Asymptotic values of peak gain G and quality factor $Q_{3 \text{ dB}}$ are both strongly increasing functions of the parameter Q_{line} . The bidirectional cochlea is not unconditionally stable, unlike the unidirectional cochlea. The maximum allowable value of Q_{line} is set by increased inter-stage reflections and the eventual onset of instability [1].

A simplified equivalent circuit of a single cochlear stage, useful for noise calculations, is derived from [1] and shown in Fig. 9. Here C_3 represents the local shunt admittance $Y(n)$, while the other components comprise the local series impedance $Z(n)$. The nodes on voltages n , $n - 1$ and m are denoted by $v(n)$, $v(n - 1)$, and $v(m)$, respectively. Nodes n and $n - 1$ are shared with adjacent stages, while node m is internal to $Z(n)$. The input impedances of the sections of transmission line basal and apical to stage n are denoted by $Z_{in-}(n - 1)$ and $Z_{in+}(n)$, respectively. We can find the values of $Z_{in-}(n - 1)$ and $Z_{in+}(n)$ iteratively, as shown in (18).

The noise generated by R_1 and R_2 is represented by the current sources i_{n1} and i_{n2} . However, R_1 is an active, negative resistor, and contains additional noise sources, such as transistors that act as current sources or sinks [1]. The effects of such noisy devices have been lumped into the additional noise current source i_{n3} . This source also models the noise produced by

any resistance in parallel with C_3 . Since $Z_{in-}(n - 1)$, $Z_{in+}(n)$ and the values of the various resistors, inductors, and capacitors in Fig. 9 are known, we can analytically find transfer impedances from the noise sources to the nodes n and $n - 1$. The transfer impedance $Z_{j,i}$ is defined the ratio of the voltage produced at node j in a network when a current is fed into node i . The two nodes may be identical, in which case $Z_{i,i}$ is known as the driving-point impedance.

By using superposition, we can find the PSD of the total noise voltage produced at node n . The result is

$$\overline{v_{ns,n}^2} = \overline{i_{n1}^2} |Z_{n,n-1} - Z_{n,n}|^2 + \overline{i_{n2}^2} |Z_{n,m} - Z_{n,n}|^2 + \overline{i_{n3}^2} |Z_{n,n}|^2. \quad (18)$$

The discrete transmission line equations shown in (16) are used to find transfer functions $TF_{j,n}(\omega)$. Each transfer function is defined as the voltage produced at node j in response to an input source placed at node n , divided by the voltage at node n . Signal transfer functions consist of the subset for which $n = 1$, i.e., the source is placed at the first node of the transmission line. When the voltage at node n is the noise $v_{ns,n}$, from (16) we have

$$\overline{v_{out,j,n}^2} = \overline{v_{ns,n}^2} [|TF_{j,n}(\omega)|^2 - |TF_{j-1,n}(\omega)|^2] \quad (19)$$

where $v_{out,j,n}$ is the output noise of the j th stage due to noise produced by the n th stage. Here the output noise of stage j is defined as the differential noise voltage between transmission line nodes j and $j - 1$. The PSD of the total noise voltage across the j th stage is the superposition of noise from all stages, i.e., for $i = 1, 2, \dots, N$, where N is the total number of stages. Thus, the PSD of the j th output of the cochlea is given by

$$\overline{v_{out,j}^2} = \sum_{n=1}^N \overline{v_{out,j,n}^2} = \sum_{n=1}^N \overline{v_{ns,n}^2} [|TF_{j,n}(\omega)|^2 - |TF_{j-1,n}(\omega)|^2] \quad (20)$$

where $v_{ns,n}$ is given by (18). The total noise voltage $v_{total,n}$ of the n th stage can be found by integrating the relevant PSD over all frequencies if the noise current PSDs $\overline{i_{n1}^2}$, $\overline{i_{n2}^2}$, and $\overline{i_{n3}^2}$ are known. Ignoring flicker noise, each noise current PSD is white and equal to $4kT\gamma g_m$. Here g_m is the small-signal transconductance of a transistor, or the conductance of a resistor, while γ is an excess noise factor that accounts for the presence of multiple noisy devices. Fig. 10 shows the results of calculating $v_{total,n}$ for a maximum operating frequency of $\omega(0) = 2\pi \times 10$ GHz and various values of d . The total number of cochlear stages was $N = 3N_{nat}$. The figure also shows that the results of a SPICE simulation of an on-chip bidirectional cochlea, which had $d \approx 0.15$, agree with theoretical calculations. We see that, as in the unidirectional case shown in Fig. 5, the noise in the bidirectional cochlea is non-monotonic with n . However, the decrease in noise for large values of n is steeper for the bidirectional cochlea than the unidirectional cochlea.

III. LOWPASS-TO-BANDPASS TRANSFORMATIONS IN THE RF COCHLEA

The RF cochlea is a rapid spectrum analyzer that provides a fast big-picture view of a broad frequency range using a low-pass representation. We have analyzed some of its properties in Section II including noise, dynamic range, and frequency resolution. Frequently, in a spectrum analyzer, a lowpass representation is not convenient since the analyzed frequency bins have

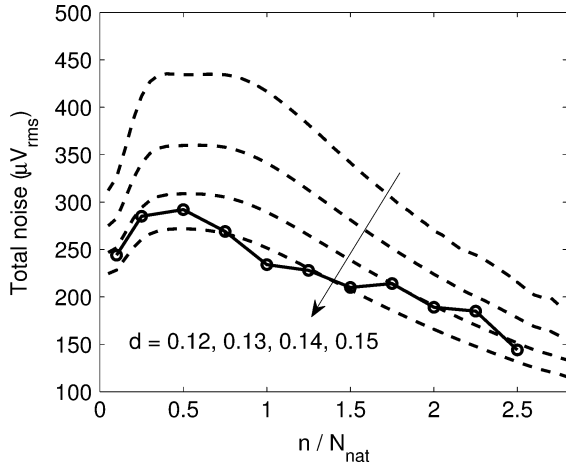


Fig. 10. Output noise voltages of the bidirectional cochlea at various values of d . Theoretical values are shown as dashed lines, while the solid line is the result of a SPICE simulation of the integrated circuit implementation. Parameters used to draw the theoretical curves were $\mu = 0.76$, $Q = 3.8$, $Q_{\text{line}} = 0.05$, and $N_{\text{nat}} = 20$.

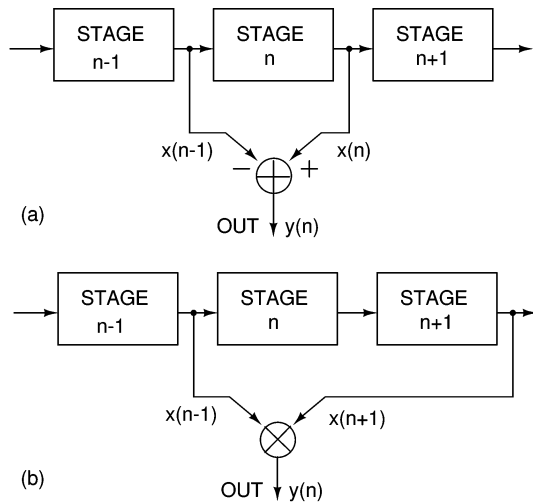


Fig. 11. Lowpass to bandpass transformations for the RF cochlea: (a) a linear differencing scheme and (b) a nonlinear center-surround scheme inspired by the action of coincidence-detecting cells in the biological cochlear nucleus.

overlapping content. In this section, we discuss how a lowpass cochlear representation can be converted into a bandpass representation using both linear and nonlinear methods. These lowpass-to-bandpass transformations are applicable to both unidirectional and bidirectional cochleas.

A simple method for implementing a lowpass-to-bandpass transformation in the cochlea is to perform discrete spatial differentiation as shown in Fig. 11(a). Thus, if we set $TF_n = G_n e^{j\phi_n}$ and perform some algebra, we obtain

$$\begin{aligned} TF_{n,\text{diff}} &\equiv TF_{n+1} - TF_n \\ &\approx \frac{\partial TF_n}{\partial n} \\ &= TF_n \left[\frac{1}{G_n} \frac{dG_n}{dn} + j \frac{d\phi_n}{dn} \right]. \end{aligned} \quad (21)$$

Equation (2) shows that frequency and spatial derivatives in the cochlea are proportional to each other. Thus spatial differ-

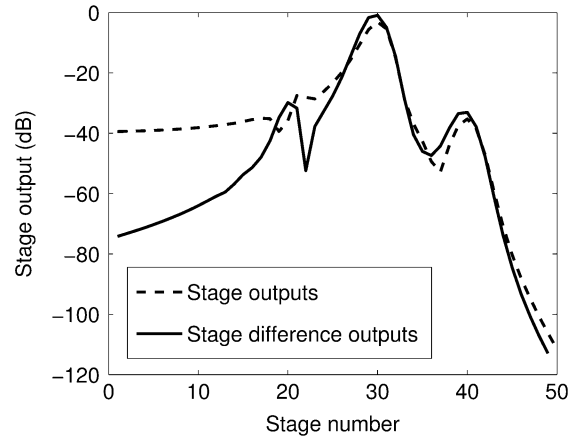


Fig. 12. Simulated outputs (spatial response using a linear differencing scheme) of the unidirectional RF cochlea to three input tones spaced one octave apart. The middle tone is 30 dB larger than the other two.

entiation converts a lowpass cochlear frequency response into a bandpass response. Some simulated spatial responses are shown in Fig. 12 for the unidirectional cochlea when three input tones are present. The small, highest-frequency tone is normally almost drowned out; however it is resolved clearly when stage differences are used as outputs.

In addition to linear differencing operations on the RF cochlea outputs, nonlinear techniques can also be used to implement a lowpass-to-bandpass transformation. These techniques can further improve the frequency selectivity of the system by utilizing spatiotemporal correlations between the outputs of cochlear stages [33]. One simple center-surround coincidence-detection strategy, inspired by the action of coincidence-detecting “octopus” cells in the biological cochlear nucleus (CN), is shown in Fig. 11(b) [34]. The CN receives inputs from the cochlea via the auditory nerve, and projects to structures higher up in the auditory pathway, such as the superior olivary complex and the inferior colliculus. Its main task is to extract simple spatio-temporal features, such as the onset of a sound and the direction of frequency sweeps (chirps).

In our strategy, outputs of cochlear stages two positions apart in the cascade are multiplied together to produce the bandpass output. The multipliers act as analog correlators that detect phase-coincidence, thus combining phase and magnitude information present in the cochlear outputs. The average output of a multiplier driven by two sinusoidal inputs with a phase difference θ between them is given by

$$\bar{y} = \overline{x \sin(\omega t) \times x \sin(\omega t + \theta)} = \frac{x^2}{2} \cos(\theta) \quad (22)$$

where the averaging operation is carried out over a complete input cycle. We see that the average output magnitude is maximal, and equal to the average power of each input signal, when $|\cos(\theta)| = 1$, i.e., when $\theta = \{0, \pi\}$. The output is zero when the inputs are in quadrature, i.e., $\theta = \{\pi/2, 3\pi/2\}$.

In our case θ corresponds to the sum of the phase differences across two adjacent cochlear stages. At frequencies much lower than the local center frequency θ is small and the output magnitude is maximal. The phase shift θ increases with frequency,

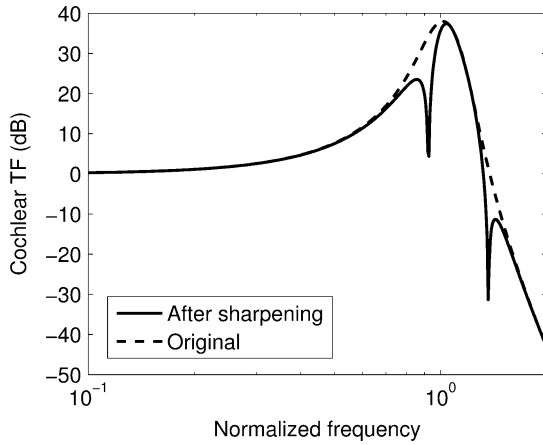


Fig. 13. Simulated unidirectional cochlea transfer functions at different spatial locations before and after center-surround coincidence detection of the cochlear outputs.

causing the output magnitude to decrease. At frequencies just higher and lower than the center frequency $\theta = \pi/2$, resulting in nulls (zero output). Finally, at the center frequency $\theta \approx \pi$, the inputs to the multiplier are almost in phase, and the output magnitude is again maximized. Thus we expect $y(n)$ to exhibit nulls on either side of the peak, but otherwise track the input amplitude $x(n)$. The whole scheme effectively decreases the width of the peak. In fact, it can be shown that our strategy approximately doubles the frequency resolution with no loss in timing precision. Simulation results, shown in Fig. 13, confirm our predictions.

Both strategies shown in Fig. 11 are used in biological auditory and visual systems [33]. The first, linear strategy is implemented using lateral inhibition networks that compute spatial derivatives. The outputs of lateral inhibition networks are then fed into networks of coincidence detectors. The actions of these neurons can be modeled using analog multipliers, as in our second, nonlinear strategy. We can combine the two strategies in a similar way to efficiently extract features, such as edges and peaks, from the RF spectrum.

IV. RF FOVEA

Lowpass-to-bandpass cochlear transformations create a more convenient spectral representation. However, they still do not provide enough frequency resolution to separate out finely-spaced narrowband signals such as adjacent wireless channels in the same frequency band, e.g., adjacent cell-phone or WiFi channels. To solve this problem, we were inspired by a solution employed by the eye and the visual system: The eye and the visual system first obtain a rapid “big-picture” view of the visual scene at moderate spatial resolution. The eye then moves to “foveate”, that is to attend to a narrow portion of the image in its central region known as the fovea, where it has much higher spatial resolution. Physical motion in the visual system is analogous to frequency motion in the auditory system, and can be implemented using mixing or heterodyning. This analogy leads to the creation of an “RF fovea”, which we describe in this section. Intuitively, the RF cochlea provides

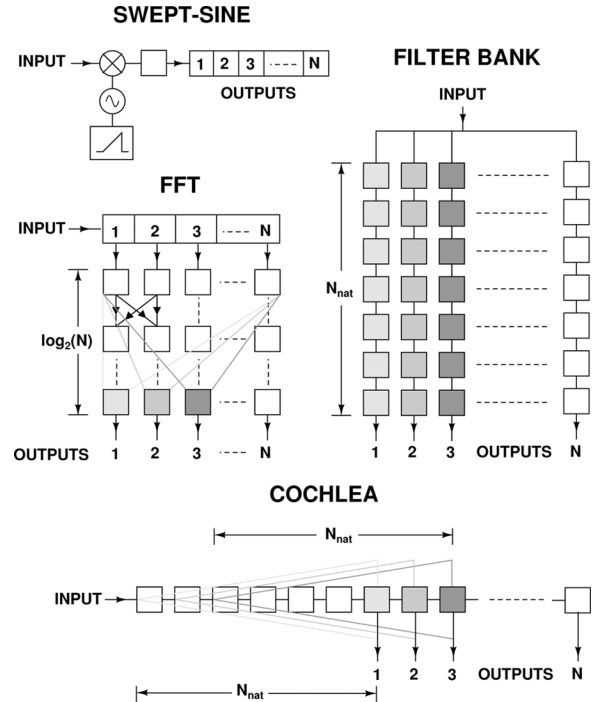


Fig. 14. Graphical illustration of several common spectrum-analysis algorithms.

the fast “big-picture” view of the entire RF spectrum at moderate frequency resolution while the RF fovea “zooms” into more narrow regions of the RF spectrum at higher frequency resolution.

The RF fovea employs successive stages of mixing, which are very similar to the successive subranging stages employed by certain analog-to-digital converters (ADCs). Thus, we shall first describe similarities between ADCs and frequency estimators. Then, we shall describe the construction of architectures for an RF fovea.

A. ADCs and Frequency Estimators

We define *frequency estimators* as systems that estimate the instantaneous frequencies present in a signal. Frequency estimation is a well-known and studied example of the general problem of estimating parameters of a signal in the presence of noise [35], [36].

Spectrum analyzers are frequency estimators that can, in addition to estimating the frequencies present in a signal, also determine their amplitude, i.e., the power spectral density, and/or phase. They may also be viewed as a parallel, broadband generalization of narrowband radios that estimate amplitude or phase of a narrowband signal centered around a single carrier frequency [37].

Several common spectrum-analysis algorithms are graphically illustrated in Fig. 14. Their performance can be conveniently compared using two metrics: hardware cost and analysis time. Both these quantities are functions of N , the number of frequency “bins” that can be estimated by the algorithm. The cochlear algorithm has been shown to be extremely efficient: Both hardware cost and analysis time scale linearly with N , i.e., as $O(N)$. The cochlear algorithm, which reuses simple

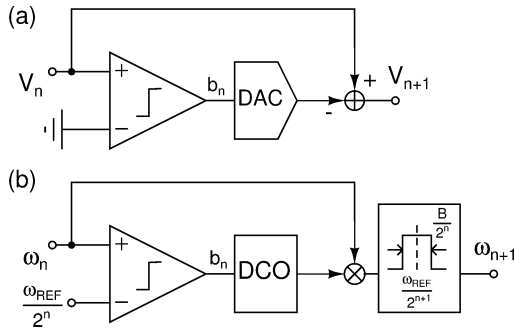


Fig. 15. Single successive-subranging stage for (a) an ADC and (b) a frequency estimator.

low- Q filters with overlapping frequency responses many times in order to synthesize high-resolution transfer functions, has low settling time and hardware complexity compared to swept-sine analyzers, FFTs, and parallel filter banks [1].

We have found interesting analogies between frequency estimators and ADCs. The problems that frequency estimators and ADCs solve are similar, but in different signal domains. Frequency estimation with a finite amount of precision may be viewed as the process by which we determine into which of N outputs, or frequency bins, to place an analog frequency variable that may be located anywhere within a bandwidth B . Similarly, ADCs find the digital value, or bin, within which an unknown analog amplitude (voltage or current) belongs. As a result, analogous algorithms work in the two cases. However, a signal can have many instantaneous frequencies of interest, but only one instantaneous amplitude. A more accurate analogy is therefore to view frequency estimators as the frequency-domain analogs of multiple parallel ADCs. Each ADC corresponds to estimating the frequency value (and possibly the power content) of one output bin.

The basic operation of subtraction in the amplitude domain is extensively used by ADCs and has a natural frequency-domain analog through multiplication (mixing or heterodyning). A low-pass or bandpass filter at the output of the mixer can be used to select the difference frequency component, while rejecting the unwanted sum frequency component, as in direct conversion receivers or traditional super-heterodyne receivers, respectively. In this paper, we shall assume that all signals are real. However, the architecture that we describe can be extended to the quadrature signal domain as well.

A simplified view of the n th stage of a generic successive-subranging ADC is shown in Fig. 15(a). The input V_n is limited to the range $\pm V_{FS}/2^n$, where V_{FS} is the initial, or full-scale range. The output of the comparator is $b_n \in [-1, 1]$, the next output bit. It is converted into an analog signal by a one-bit digital-analog converter (DAC) whose output voltage is $b_n V_{FS}/2^{n+1}$. The DAC output is subtracted from the input V_n to extract a residue V_{n+1} that spans half the range of V_n and is the input to the next conversion stage. In general the one-bit quantizer and DAC can be replaced by multiple-bit versions, and the stages pipelined or reused during the conversion cycle.

We can begin to map the ADC of Fig. 15(a) to a frequency estimator in Fig. 15(b) by replacing the full-scale voltage range

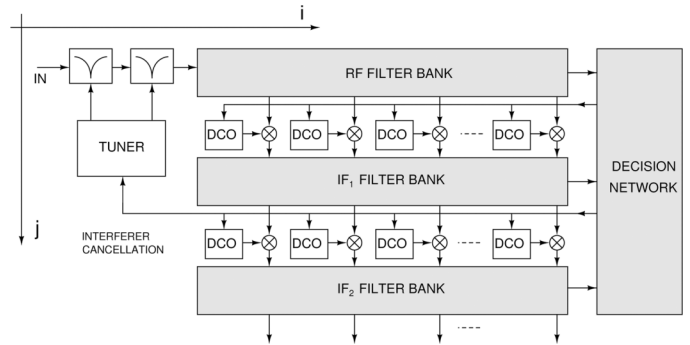


Fig. 16. Cascaded super-heterodyne architecture: Just as in successive subranging ADCs, the frequency information is focused on or estimated to higher and higher precision in each stage of the architecture. In this example, the architecture is used in a feedback-loop configuration such that interfering tones can be estimated and then canceled out using tunable notch filters.

with an initial frequency range B and by replacing the voltage comparator with a frequency comparator. However, frequency estimation requires estimation of bandpass input signals, analogous to an ADC that requires signals within a certain positive voltage range to be digitized. Hence, at each recursive conversion stage in a frequency estimator, the offset or “common-mode” frequency used to compare the input frequency against and the bandwidth of the conversion decrease by a factor of 2. Thus, in Fig. 15(b), the DAC of Fig. 15(a) is replaced with a digitally-controlled oscillator (DCO) whose output frequency is $\omega_{DCO} = (\omega_{REF} + b_n B)/2^{n+1}$ and the subtractor of Fig. 15(a) is replaced by a mixer and bandpass filter as shown.

The frequency comparator can be built by converting the input and reference frequencies into voltages, and then using a voltage comparator. Frequency-to-voltage converters can be created using tuned circuits (discriminators) or phase-locked loops. The converters should be matched; otherwise the comparator will have a static offset. In addition, the functional relationship between input frequency and output voltage should be monotonic (it need not be linear).

B. Cochlear Heterodyning

The RF spectrum is sparse over large bandwidths, i.e., dominated by a limited number of relatively narrowband signals which cumulatively occupy only a small fraction of the total bandwidth. Some of these signals are of interest to the user, while others, referred to as interferers, are not. In practical situations interferers can be much larger than desirable signals. A system capable of simultaneously observing several narrow, arbitrarily placed frequency bands can therefore capture essentially all significant information present within the entire RF spectrum. An architecture that uses successive subranging in the frequency domain to solve this problem is shown in Fig. 16. This *cascaded super-heterodyne architecture* decides where observed regions of the spectrum should be located by using filter banks to decompose spectra into multiple channels.

The filter banks in Fig. 16 have bandpass transfer functions with center frequencies and bandwidths that scale arbitrarily with horizontal position i . The outputs at each position are analyzed by a cascade of successive subranging stages, each of which is similar to that shown in Fig. 15(b). We may identify

$\omega_{\text{REF}}/2^n$ and $B/2^n$ with the local center frequency and bandwidth, respectively. The latter can remain constant with i , or scale with center frequency to generate constant- Q characteristics. Here $Q = \omega_{\text{REF}}/B$ is the quality factor. In addition, for a fixed value of i the center frequencies and bandwidths change with vertical position j , usually decreasing by a constant factor α at each row. The bandwidth of the i th filter in the j th bank, which we denote by $B_{j,i}$, is given by

$$B_{j,i} = \frac{B_{j-1,i}}{\alpha} = \frac{B_{0,i}}{\alpha^m}. \quad (23)$$

A common value for α is 2, in which case the decision network consists of one-bit quantizers, i.e., frequency comparators that determine which half of the input bandwidth $B_{j-1,i}$ should be included in the output bandwidth $B_{j,i}$. In general we must use quantizers with $\log_2(\alpha)$ bits. Thus, the output of each successive filter bank allows the user to determine the frequencies of N signals of interest in parallel with exponentially-increasing precision, where N is the number of filters in each filter bank. The outputs of the decision network can be viewed as the results of N parallel frequency-to-digital conversions (analogous to analog-to-digital conversions). The architecture is also efficient in terms of hardware requirements because it only needs M filter banks and $M \times N$ mixers to increase precision by a factor of α^M . Finally, it retains amplitude information about each output signal and thus acts as a spectrum analyzer as well as a frequency estimator.

The ultimate precision of successive subbranching architectures for frequency estimation is limited by three factors: quantization noise, random errors (caused by device mismatch, thermal noise, and flicker noise), and interference from signals in adjacent channels. The first two factors also affect the performance of ADCs, while the third is unique to frequency estimation. Such interference is due to overlap between the transfer functions of the initial filter bank, and can be reduced by using higher-order filters with larger roll-off slopes and stopband rejections. The main function of subsequent filter banks is to further reduce the relative amount of interference through additional filtering. They can be omitted, or their order reduced, if interference levels are low enough.

One application of the cascaded super-heterodyne structure is to accurately estimate the frequencies and bandwidths of unwanted interferers. The interferers can then be canceled out, allowing weak signals to be detected in their presence. A simple interference-cancellation strategy that uses tunable notch filters is shown in Fig. 16. This scheme was simulated using MATLAB; the results are shown in Fig. 17. Two tones one octave apart are fed into the structure, with one being 80 dB larger in amplitude than the other. The curve labeled “original” shows the results without interferer cancellation: The smaller tone is invisible. The lower curve shows filter bank outputs after the large interfering tone has been estimated and notched out: The small tone is now clearly resolved.

Cochlea-like traveling-wave structures consist of transmission lines or filter cascades with maximal-response frequencies that decrease exponentially with position. They are an efficient way to synthesize high-order, constant- Q filter banks because of extensive hardware reuse: Each output is filtered by multiple

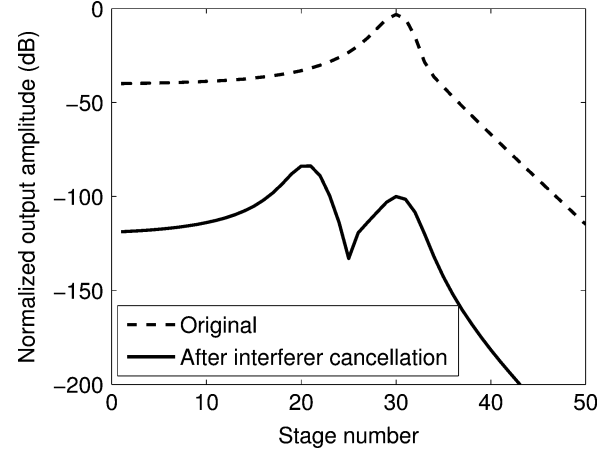


Fig. 17. Simulated performance of the cascaded super-heterodyne architecture. The plots show outputs of the final filter bank before and after a large interfering signal was estimated and canceled out.

stages, each of which simultaneously produces its own output. Unlike other spectrum analysis algorithms, such as the FFT or simple filter banks, each cochlear stage is used multiple times to synthesize output transfer functions with large roll-off slopes, and consequently, high frequency selectivities. As a result, the hardware and power costs of the cochlea are much lower than comparable algorithms [1].

An added advantage of constant- Q frequency decomposition is that it is better matched to the real RF environment. The RF spectrum today is divided into licensed and unlicensed bands that roughly follow constant- Q characteristics, i.e., have bandwidths that are approximately proportional to center frequencies. For example, the Q of the unlicensed industrial-scientific-military (ISM) frequency bands allocated internationally only varies by about a factor of ten as the center frequencies vary by three orders of magnitude. It is thus advantageous to replace the filter banks in Fig. 16 with cochlea-like traveling-wave structures. We refer to this special case, which is shown in Fig. 18, as *cochlear heterodyning*. It can be used to create an *RF fovea* just as in Fig. 16, i.e., an architecture where an RF cochlea rapidly obtains an overview of the entire spectrum, and then adaptively “locks in” or “foveates” on desired regions in that spectrum.

As in Fig. 16, center frequencies and bandwidths in Fig. 18 decrease exponentially with vertical position j . The center frequency of the j th output from the i th cochlea is proportional to its bandwidth (constant- Q behavior) and is given by

$$f_{i,j} = f_{0,0} 2^{-i/N_{\text{Oct}}} \alpha^{-j} = f_{0,0} e^{-[i \ln(2)/N_{\text{nat}} + j \ln(\alpha)]} \quad (24)$$

where each cochlea is assumed to produce N_{Oct} outputs per octave, and α is the constant that represents the exponentially increasing precision as we move from RF to lower-frequency IF stages in Fig. 18. The outputs of each cochlea are downconverted using mixers. Each downconversion step selects $1/\alpha$ of the output bandwidth of the previous cochlear output. For example, if $\alpha = 2$, one-bit decisions determine which half gets selected by changing the DCO frequency f_{ji} fed into the mixer between $f_{ji,0}$ and $f_{ji,1}$, where $1 \leq j \leq M$ and $1 \leq i \leq N$ are the cochlear and output indices, respectively.

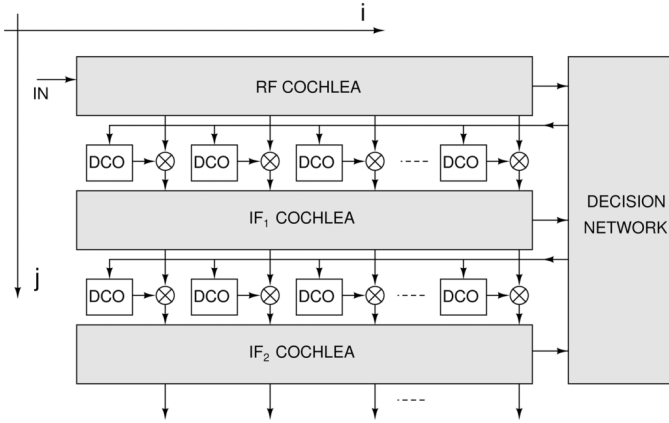


Fig. 18. Cochlear heterodyning, a broadband generalization of narrowband heterodyning where narrowband RF, IF, and baseband stages are replaced by RF, IF, and baseband cochlear stages. In this example, two IF stages are shown. In each stage, successively finer frequency resolution is obtained enabling foveation onto interesting regions of the RF spectrum.

In order to reduce hardware costs we can downconvert only those outputs that appear significant, such as local maxima (peaks). The output of the mixer at location i can be directly fed into the i th stage of the next cochlea, as shown in Fig. 18. Alternatively, all outputs of the j th cochlea can be added and fed into the first stage of the $j + 1$ th cochlea.

The total acquisition time of the RF fovea is determined by the sum of the settling times of all the cochleas. It is dominated by the final cochlea since it analyzes the lowest input frequencies, and is given by

$$T_{\text{acq}} \approx \frac{\gamma N_{\text{oct}}}{f_{M,N}} \left(\frac{\alpha}{\alpha - 1} \right) \quad (25)$$

where γ is a constant, of order unity, that depends on the cochlear model, $f_{M,N} = f_{0,N}/\alpha^M$ is the lowest frequency analyzed by the final cochlea, there are a total of M cochleas and $f_{0,N}$ is the lowest frequency analyzed by the first (input) cochlea. The hardware complexity of the system is modest: It scales as $O(NM)$, where N is given by

$$N = N_{\text{oct}}(\log_2(\beta) + 1). \quad (26)$$

This formula assumes (conservatively) that the first N_{oct} stages of each cochlea are necessary for building up the transfer functions and do not produce useful outputs. Also, $\beta = f_{j,N_{\text{oct}}}/f_{j,N}$ is defined as the ratio of maximum and minimum frequencies analyzed by each cochlea.

Fig. 19 shows the simulated frequency estimation error of the RF fovea for various values of M . We generated this plot by feeding artificial noiseless spectra into the system. Each spectrum consisted of N_f tones with random frequencies and amplitudes, uniformly distributed in log-frequency and log-amplitude space, respectively. We used the unidirectional cochlear model described in Section II, and set $N_{\text{oct}} = 10$ and $\alpha = 2$ for the simulations.

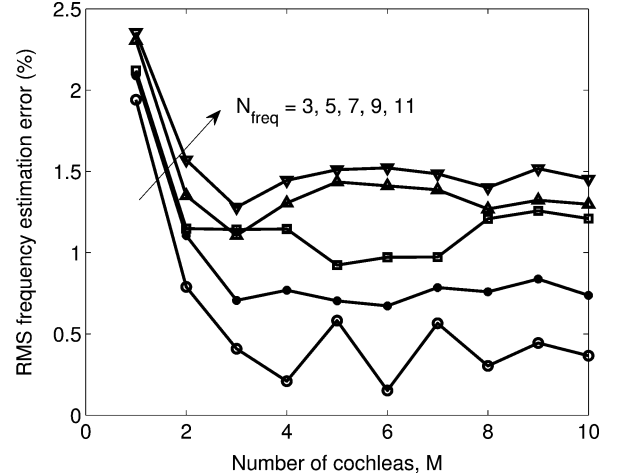


Fig. 19. Frequency estimation error of the RF fovea to inputs consisting of N_f tones with random frequencies and amplitudes as a function of M , the total number of cochleas, obtained from MATLAB simulations.

At the end of the successive-subranging procedure, amplitude peak picking was used to map spatial positions to frequencies. Using (14), the resultant fractional estimation error is given by

$$\sigma_{\text{final}} = \frac{\sigma_0}{\alpha^M}. \quad (27)$$

We used $N_{\text{oct}} = 10$ in our simulations, corresponding to $\sigma_0 = 2.1\%$. Our simulations confirm that the frequency-estimation error decreases with M , as predicted by (27). However, the error asymptotes to a fixed value for large values of M . This asymptotic error increases monotonically with N_f , indicating that the precision of our algorithm is being limited by interference from other parts of the spectrum that have not been notched out as in Fig. 16. Besides notching, we can also increase N_{nat} , which is proportional to the high-frequency rolloff slopes of the cochlear transfer functions, to reduce such interference. The result is improved precision at the cost of increased hardware and power consumption. For a given amount of interference the final estimation error can also be improved by using a more sophisticated algorithm than simple amplitude peak picking. For example, phase information can be used, as suggested by Fig. 8.

In the presence of noise the frequency estimation error increases because the successive-subranging networks make wrong decisions. When simple amplitude detection is used in the presence of Gaussian noise, the error probability of each decision is given by

$$P_e(m) = \frac{1}{2} \text{erfc} \left(\frac{\text{SNR}(m)}{\sqrt{2}} \right) \quad (28)$$

where $\text{SNR}(m)$ is the signal-to-noise ratio in the m th cochlea. The overall fractional estimation error can be estimated by summing up the error probabilities in each of the M cochleas.

An erroneous decision can occur at any of the M successive-subranging operations. Once an error occurs the algorithm converges to a random answer set by noise. Assuming Gaussian

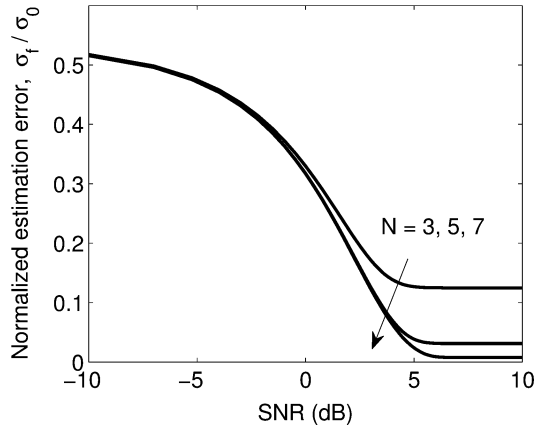


Fig. 20. Normalized RMS frequency estimation error of the RF fovea as a function of SNR and the number of cochleas M .

noise, an erroneous decision in the m th cochlea results in a fractional frequency estimation error that is a zero-mean Gaussian random variable with variance equal to

$$\sigma_e^2(m) = \frac{2\sigma_0^2}{\alpha^{2m}}. \quad (29)$$

This formula can be easily derived from (14) by noting that the frequency resolution increases by a factor of α after each cochlea. The factor of 2 arises as follows. When an error occurs, the desired answer, i.e., the correct frequency, remains unknown, and also acts as a random variable. Therefore, the difference between the incorrect and correct answers, which is the estimation error, is a random variable with double the variance of the final incorrect answer.

The probability that the first error occurs at the m th decision is given by

$$w_e(m) = P_e(m) \prod_{n=1}^{m-1} (1 - P_e(n)). \quad (30)$$

When no errors occur the algorithm converges to the correct answer, resulting in an estimation error that has a variance of σ_{final}^2 , as shown in (27). The probability that no errors occur is given by

$$w_{ne} = \prod_{n=1}^M (1 - P_e(n)). \quad (31)$$

The final estimation error is a random variable that is produced by a Gaussian mixture process with $M + 1$ components. The first M components have variances given by (29) and probabilities, or weights, given by (30). The last component has a variance of σ_{final}^2 and a probability given by (31). Thus the final estimation error is also Gaussian, and has a variance given by

$$\sigma_f^2 = \sum_{m=1}^M w_e(m) \sigma_e^2(m) + w_{ne} \sigma_{\text{final}}^2. \quad (32)$$

The first term in (32) corresponds to the case when at least one of the decision networks makes an error. The second term corresponds to the case when all the decisions are correct. It is the product of two terms: the probability that all the decisions are correct, and the final estimation error σ_{final} . Equation (32) is plotted in Fig. 20 for the case when SNR has the same

value in all M cochleas. The estimation error decreases as SNR increases, but asymptotes to σ_{final} for large-enough SNR. The asymptotic SNR is quite modest, as shown in Fig. 20.

V. CONCLUSION

We have analyzed the noise, dynamic range, and frequency resolution of cochlea-like traveling-wave structures such as the RF cochlea. Our analysis showed that frequency resolution is significantly improved by utilizing phase information present in the cochlear outputs rather than utilizing amplitude information alone. Thus, a cochlear-nucleus-inspired lowpass-to-bandpass center-surround coincidence detector operating on the cochlear outputs improves frequency resolution without compromising temporal resolution. To further improve frequency resolution, we also described an eye-inspired *RF fovea*. The RF fovea is a successive-subranging receiver architecture that allows users to focus on selected narrowband portions of a broadband input spectrum. The RF fovea is built with a *cochlear heterodyning* architecture, which generalizes the idea of heterodyning from the narrowband domain to the broadband domain. It may be useful as a dynamic spectrum-sensing front end for software-defined and cognitive radio receivers. For example, the dynamic range of such receivers can be increased by detecting strong unwanted signals (interferers) and adaptively removing them in a feedback loop. Our cochlear models use gain control to increase dynamic range, i.e., they amplify strong signals less than weak signals. The gain-control process causes two-tone suppression: Strong tones reduce the gain of weaker signals at nearby frequencies, thus enhancing spectral contrast [38]. As a result our system can detect interferers in noisy environments more reliably than purely linear schemes.

REFERENCES

- [1] S. Mandal, S. Zhak, and R. Sarpeshkar, "A bio-inspired active radio-frequency silicon cochlea," *IEEE J. Solid-State Circuits*, vol. 44, no. 6, pp. 1814–1828, Jun. 2009.
- [2] R. Sarpeshkar, R. F. Lyon, and C. Mead, "A low-power wide-dynamic-range analog VLSI cochlea," *Analog Integr. Circuits Signal Process.*, vol. 16, pp. 245–274, Aug. 1998.
- [3] A. Katsiamis, E. Drakakis, and R. Lyon, "A biomimetic, 4.5 μW , 120+ dB, log-domain cochlea channel with AGC," *IEEE J. Solid-State Circuits*, vol. 44, no. 3, pp. 1006–1022, Mar. 2009.
- [4] R. Sarpeshkar, *Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-Inspired Systems*, 1st ed. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [5] L. Watts, "Real-time, high-resolution simulation of the auditory pathway, with application to cell-phone noise reduction," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2010, pp. 3821–3824.
- [6] J. Mitola, III and G. Q. Maguire, Jr., "Cognitive radio: Making software radios more personal," *IEEE Personal Commun.*, vol. 6, no. 4, pp. 13–18, Aug. 1999.
- [7] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [8] S. Haykin, J. H. Reed, G. Y. Li, and M. Shafi, Eds., *Special Issue on Cognitive Radio, ser. Proc. IEEE*, vol. 97, no. 4–5, Apr.–May 2009.
- [9] J. Mitola, III, "Software radios: Survey, critical evaluation and future directions," *IEEE Aerosp. Electron. Syst. Mag.*, vol. 8, no. 4, pp. 25–36, Apr. 1993.
- [10] J. Mitola, III, "The software radio architecture," *IEEE Commun. Mag.*, vol. 33, no. 5, pp. 26–38, May 1995.
- [11] F. Rivet, Y. Deval, J.-B. Begueret, D. Dallet, P. Cathelin, and D. Belot, "A disruptive receiver architecture dedicated to software-defined radio," *IEEE Trans. Circuits Syst. II, Exp. Papers*, vol. 55, no. 4, pp. 344–348, Apr. 2008.

- [12] P. K. Prakasam, M. Kulkarni, X. Chen, Z. Yu, S. Hoyos, J. Silva-Martinez, and E. Sánchez-Sinencio, "Applications of multipath transform-domain charge-sampling wide-band receivers," *IEEE Trans. Circuits Syst. II, Exp. Papers*, vol. 55, no. 4, pp. 309–313, Apr. 2008.
- [13] B. Razavi, "Cognitive radio design challenges and techniques," *IEEE J. Solid-State Circuits*, vol. 45, no. 8, pp. 1542–1553, Aug. 2010.
- [14] R. Negra, A. Sadeve, S. Bensmida, and F. Ghannouchi, "Concurrent dual-band class-F load coupling network for applications at 1.7 and 2.14 GHz," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, no. 3, pp. 259–263, Mar. 2008.
- [15] K. L. Melde, H.-J. Park, H.-H. Yeh, B. Fankem, Z. Zhou, and W. R. Eisenstadt, "Software defined match control circuit integrated with a planar inverted F antenna," *IEEE Trans. Ant. Propag.*, vol. 58, no. 12, pp. 3884–3890, Dec. 2010.
- [16] Z. Ru, E. Klumperink, and B. Nauta, "Discrete-time mixing receiver architecture for RF-sampling software-defined radio," *IEEE J. Solid-State Circuits*, vol. 45, no. 9, pp. 1732–1745, Sep. 2010.
- [17] C. Galbraith, R. White, L. Cheng, K. Grosh, and G. Rebeiz, "Cochlea-based RF channelizing filters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 4, pp. 969–979, Apr. 2008.
- [18] R. Mahesh and A. Vinod, "Reconfigurable frequency response masking filters for software radio channelization," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 55, no. 3, pp. 274–278, Mar. 2008.
- [19] H. Shin and Y. Kim, "A CMOS active-RC low-pass filter with simultaneously tunable high- and low-cutoff frequencies for IEEE 802.22 applications," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 2, pp. 85–89, Feb. 2010.
- [20] S. Osmany, F. Herzel, and J. Scheytt, "An integrated 0.6–4.6 GHz, 5–7 GHz, 10–14 GHz, and 20–28 GHz frequency synthesizer for software-defined radio applications," *IEEE J. Solid-State Circuits*, vol. 45, no. 9, pp. 1657–1668, Sep. 2010.
- [21] J. Park, T. Song, J. Hur, S. M. Lee, J. Choi, K. Kim, K. Lim, C.-H. Lee, H. Kim, and J. Laskar, "A fully integrated UHF-band CMOS receiver with multi-resolution spectrum sensing (MRSS) functionality for IEEE 802.22 cognitive radio applications," *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp. 258–268, Jan. 2009.
- [22] C.-J. Li, F.-K. Wang, T.-S. Horng, and K.-C. Peng, "A novel RF sensing circuit using injection locking and frequency demodulation for cognitive radio applications," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 12, pp. 3143–3152, Dec. 2009.
- [23] T. Song, J. Park, S. M. Lee, J. Choi, K. Kim, C.-H. Lee, K. Lim, and J. Laskar, "A 122-mW low-power multiresolution spectrum-sensing IC with self-deactivated partial swing techniques," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 57, no. 3, pp. 188–192, Mar. 2010.
- [24] C. Lelandais-Perrault, T. Petrescu, D. Poulton, P. Duhamel, and J. Oksman, "Wideband, bandpass, and versatile hybrid filter bank A/D conversion for software radio," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 56, no. 8, pp. 1772–1782, Aug. 2009.
- [25] M. Mishra and A. Sahai, "How much white space is there?," EECS Dept., Univ. California, Berkeley, Jan. 2009. [Online]. Available: <http://www.eecs.berkeley.edu/Pubs/TechRpts/2009/EECS-2009-3.html>
- [26] S. Mandal and R. Sarpeshkar, "A cochlear heterodyning architecture for an RF fovea," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, 2010, pp. 3825–3828.
- [27] S. Mandal, "Collective analog bioelectronic computation," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, Jun. 2009.
- [28] G. Zweig, "Finding the impedance of the organ of Corti," *J. Acoustical Soc. Amer.*, vol. 89, no. 3, pp. 1229–1254, Mar. 1991.
- [29] S. M. Zhak, "Modeling and design of an active silicon cochlea," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Massachusetts Inst. Technol., Cambridge, Sep. 2008.
- [30] J. O. Pickles, *An Introduction to the Physiology of Hearing*, 2nd ed. London, U.K.: Academic Press, 1988.
- [31] O. Ghitza, "Robustness against noise: The role of timing-synchrony measurement," in *Proc. IEEE Int. Acoust., Speech, Signal Process. Conf. (ICASSP)*, 1987, vol. 12, pp. 2372–2375.
- [32] M. L. Fowler and J. A. Johnson, "Extending the threshold and frequency range for phase-based frequency estimation," *IEEE Trans. Signal Process.*, vol. 47, no. 10, pp. 2857–2863, Oct. 1999.
- [33] S. Shamma, "On the role of space and time in auditory processing," *Trends in Cognitive Sci.*, vol. 5, no. 8, pp. 340–348, Aug. 2001.
- [34] 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 Congr. Otorhinolaryngology, Head, Neck Surgery, Sydney, Australia, Mar. 1997.
- [35] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *IEEE Trans. Inf. Theory*, vol. 20, no. 5, pp. 591–598, Sep. 1974.
- [36] J. D. Klein, "Fast algorithms for single frequency estimation," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1762–1770, May 2006.
- [37] S. Mandal, S. Zhak, and R. Sarpeshkar, "Architectures for universal or software radio," U.S. Patent Appl. 11/958 990 filed Dec. 18, 2007. Published US2008/0240301 A1, Oct. 2008.
- [38] L. Turicchia and R. Sarpeshkar, "A bio-inspired companding strategy for spectral enhancement," *IEEE Trans. Speech Audio Process.*, vol. 13, no. 2, pp. 243–253, Feb. 2005.



Soumyajit Mandal (M'09) received the B.Tech. degree from the Indian Institute of Technology, Kharagpur, India, in 2002, and the S.M. and Ph.D. degrees in electrical engineering from the Massachusetts Institute of Technology (MIT), Cambridge, in 2004 and 2009, respectively.

He is currently a postdoctoral research scientist at Schlumberger-Doll Research, Cambridge, MA. His research interests include analog and biological computation, nonlinear dynamics, nuclear magnetic resonance, low-power analog and RF circuit design,

and antennas.

Dr. Mandal was a recipient of the President of India Gold Medal in 2002. In 2009 he delivered the annual MIT Microsystems Technology Laboratories (MTL) Doctoral Dissertation Seminar in recognition of outstanding research of interest to a broad audience.



Rahul Sarpeshkar (M'97–SM'07) received the Bachelor's degrees in electrical engineering and physics from the Massachusetts Institute of Technology (MIT), Cambridge and the Ph.D. degree from California Institute of Technology (CalTech), Pasadena.

After completing the Ph.D. degree, he joined Bell Labs as a member of technical staff in the Department of Biological Computation within its Physics Division. Since 1999, he has been on the faculty of Electrical Engineering and Computer Science Department, MIT, where he heads a research group on analog VLSI and biological systems. He holds over 25 patents and has authored over 100 publications including one featured on the cover of *NATURE*. He has authored the recent text, *Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-Inspired Systems*, which provides a broad and deep treatment of the fields of low-power electronics and bioelectronics.

Dr. Sarpeshkar was a recipient of several awards including the NSF Career Award, the ONR Young Investigator Award, the Packard Fellows Award, and the Indus Technovator Award for his interdisciplinary bioengineering research. He is an Associate Editor of the *IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS* and serves on the program committees of several technical conferences.