A Silicon Cochlea With Active Coupling

Bo Wen, *Member, IEEE*, and Kwabena Boahen, *Member, IEEE*

***Abstract*—We present a mixed-signal very-large-scale-integrated chip that emulates nonlinear active cochlear signal processing. Modeling the cochlea's micromechanics, including outer hair cell (OHC) electromotility, this silicon (Si) cochlea features active coupling between neighboring basilar membrane (BM) segments—a first. Neighboring BM segments, each implemented as a class AB log-domain second-order section, exchange currents representing OHC forces. This novel active-coupling architecture overcomes the major shortcomings of existing cascade and parallel filter-bank architectures, while achieving the highest number of digital outputs in an Si cochlea to date. An active-coupling architecture Si cochlea with 360 frequency channels and 2160 pulse-stream outputs occupies 10.9 mm^2 in a five-metal 1-poly 0.25-μm CMOS process. The chip's responses resemble that of a living cochlea's: Frequency responses become larger and more sharply tuned when active coupling is turned on. For instance, gain increases by 18 dB and Q_{10} increases from 0.45 to 1.14. This enhancement decreases with increasing input intensity, realizing frequency-selective automatic gain control. Further work is required to improve performance by reducing large variations from tap to tap.**

***Index Terms*—Class AB, cochlear amplifier, log-domain, neuromorphic, silicon (Si) cochlea.**

I. Silicon Cochleae

SILICON (Si) cochleae emulate cochlear processing of sound stimuli in very-large scale integrated (VLSI) systems, attempting to match the biological cochlea's sound sensitivity, frequency selectivity, and dynamic range. The effort to build artificial cochleae in Si has been largely motivated by their potential applications in hearing aids, cochlear implants, and other portable devices that demand real-time, low-power signal processing for speech recognition; these requirements favor subthreshold analog VLSI designs [1]. Furthermore, analog VLSI is amenable to cochlea-like distributed processing due to its compact computational elements, large numbers of which can be integrated in a small area of Si. However, digital outputs are easier to interface with higher level processing, whether performed by other neuromorphic chips or implemented in computer software. Thus, a mixed-mode approach, where the cochlea's analog outputs are converted to digital pulses—a function performed by the auditory nerve (AN)—is most attractive.

Manuscript received August 17, 2008; revised November 21, 2008. First published October 30, 2009; current version published November 25, 2009. This work was supported by the Packard Foundation, under Grant 99-1454. This paper was recommended by Associate Editor T. Delbruck.

B. Wen was with the Department of Bioengineering, University of Pennsylvania. She is now with the Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: bwen@mit.edu).

K. Boahen was with the University of Pennsylvania, Philadelphia, PA 19104 USA. He is now with the Department of Bioengineering, Stanford University, Stanford, CA 94305 USA (e-mail: boahen@stanford.edu).

Digital Object Identifier 10.1109/TBCAS.2009.2027127

Si cochleae take the form of a bank of low-pass or band-pass filters, with exponentially decreasing resonant frequencies, connected in cascade or in parallel. Cascaded filter banks, introduced in the first Si cochlea [2], rely on gain accumulation, with each filter's gain being small. Their major drawbacks are excessive delay and noise accumulation [3], and poor fault tolerance. Parallel filter banks require each filter to generate the desired gain and tuning by itself, falling short of the biological cochlea's frequency tuning and cutoff slopes [4]. A variation of the parallel architecture introduced by Watts [5] couples the filters together through a resistive grid that models the cochlear fluid. Although this coupled architecture emulates the cochlea more faithfully, its gain is diminished by destructive interactions [6].

Our Si cochlea aims to overcome existing architectures' shortcomings by mimicking the cochlea's micromechanics, in particular, the intricate anatomical arrangement of outer hair cells (OHCs) and other structural cells in the organ of Corti. Although it is a mystery as to how exactly OHC motile forces, discovered in mammalian cochlea more than two decades ago [7], boost the basilar membrane's (BM) vibration, cochlear microanatomy provides clues. Based on these clues, we previously proposed a novel mechanism for the cochlear amplifier—active bidirectional coupling (ABC) [8]. Here, we report a mixed-signal VLSI chip that implements ABC, the first cochlear chip that employs active behavior (i.e., negative damping [9]–[11]) instead of passive behavior (i.e., undamping [12], [13][1]).

By counteracting the coupled architecture's destructive interference, ABC promises frequency tuning comparable to human performance. The psychophysically measured auditory filter width, or critical band, is about 1/3 to 1/6 octave [14]. This bandwidth suggests a Q_{10}—center frequency divided by width 10 dB below the peak—of between 3 and 6 at the BM. In fact, Q_{10} values measured from cat AN fibers increase from 1 to above 6 from 200 to 20 kHz [15]. In comparison, the highest Q_{10} reported for the cascade and parallel architectures are 0.92 [2] and 0.42 [4], respectively. In the former, the individual filter Q must be limited to manage noise accumulation [3]; in the latter, individual filters had to achieve the desired performance on their own. By avoiding these constraints, the passively coupled architecture achieved a best-case Q_{10} of 2.34—the highest to date—despite a 25-dB gain-drop due to destructive interference [6]. Our software simulations suggested that ABC could counteract this destructive interference [8], thereby achieving performance comparable to humans. However, this proved challenging in Si: We discovered that reflections were caused by abrupt changes in BM properties (due to the transistor mismatch).

Section II presents the challenges that VLSI implementations of ABC face. Section III presents a mathematical model of

[1]The term "active" used in [13] refers to the fact that the Q is actively controlled; this differs from our use of active, which refers to actively pumping energy into the traveling wave through negative damping.

ABC, first proposed in [8]. Section IV presents the synthesis of an analog circuit that satisfies the model's equations. Section V presents a transistor-level circuit implementation. Section VI presents real-time chip responses that emulate nonlinear active cochlear behavior. Section VII discusses the impact of transistor mismatch. Section VIII concludes this paper. This paper extends the work described in [9] and [16] .

II. Implementation Challenges

While our software simulations demonstrated ABC's promise as a cochlear amplifier [8], [17], implementing it in digital or analog VLSI presents challenges. According to our simulations, a large gain (more than 60 dB) is achieved when negative damping (i.e., active amplification) occurs over many BM segments (about 60). This requirement necessitates a large number of segments per octave (about 45) if sharp tuning ($Q_{10} > 5$) is desired as well. The upshot shows that about 450 segments are needed to span the audio-frequency range (20–20 kHz or ten octaves), presenting challenges for digital and analog implementations.

As for digital VLSI, although the bit-serial technique offers implementation efficiency, it is hard-pressed to fit several hundred segments on a chip. This approach yielded 71 second-order sections in a 40 mm^2 1.2-μm-complementary metal–oxide semiconductor (CMOS) application-specific integrated circuit (ASIC) [18] and 88 sections in a Xilinx Virtex XCV1000 FPGA [19]. Extrapolating these numbers yields 409 sections in a 10 mm^2 0.25 μm-CMOS ASIC and 334 sections in a Xilinx Virtex II XC2V8000, but this does not include the fluid model nor does it include ABC.[2] Adding this functionality, which requires two multiply-accumulates for ABC and about ten for the fluid (per section), will double the complexity of the system, and halve the number of sections.

As for analog VLSI, fitting several hundred segments on a single chip is possible if small transistors and capacitors are used, but these are prone to mismatch and noise. However, given that the biological cochlea itself is built out of imprecise components, we conjectured that ABC will be robust to mismatch and noise in Si devices. For instance, noise in transistors used to model the fluid in the Si cochlea parallels Brownian motion of water molecules impinging on the basilar membrane. Our motivation for implementing ABC in analog VLSI was to explore this conjecture—if indeed ABC inherited its biological counterpart's robustness. To this end, we integrated hundreds of BM segments in a single chip, passively and actively coupled by transistors mimicking the cochlear-fluid and ABC, respectively. In addition, the chip includes Si neurons that convert analog signals, representing BM velocity, into digital pulse-streams, representing AN fibers' spike trains.

III. Nonlinear Active Cochlear Model

The cochlea actively amplifies acoustic signals as it performs spectral analysis. Incoming sound moves the oval window (stapes) at the cochlea's base, which, in turn, sets the cochlear fluid in motion [Fig. 1(a)]. The fluid interacts with the BM, the cochlea's main vibrating organ, forming a traveling wave that propagates toward the cochlea's apex. From the base to the apex, BM transverse fibers increase in width and decrease

[2]The coupling in [18] is between automatic-gain-control (AGC) filters, not between BM segments.

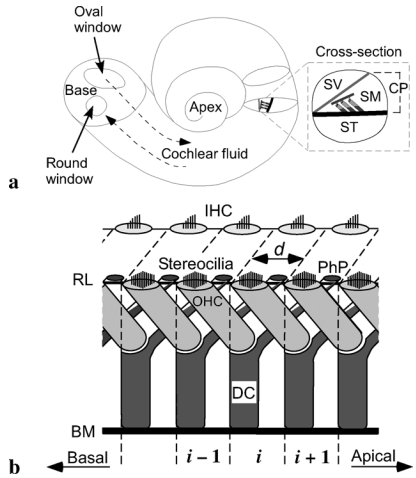


Fig. 1. Cochlea. (a) Cutaway showing cochlear ducts (adapted from [20]), comprising (inset) the scala vestibuli (SV), scala media (SM), and scala tympani (ST). The cochlear partition (CP) separates the two perilymphatic scalae. (b) Longitudinal view of the CP (adapted from [7], [21]–[23]). Outer hair cells (OHCS) tilt toward the base while Deiters' cells' (DC) phalangeal processes (PhP) tilt toward the apex; their bases rest on the basilar membrane (BM) and their tips form the reticular lamina (RL). d is the tilt distance.

in thickness, resulting in an exponential decrease in stiffness, which gives rise to the passive frequency tuning of the cochlea. BM vibration is actively enhanced by OHC electromotile forces, resulting in the cochlea's exquisite sound sensitivity, frequency discriminability, and nonlinearity.

Assuming it is incompressible, the fluid's motion can be described by a velocity potential ϕ that satisfies $\nabla^2\phi(x,y,t)=0$, where ∇^2 is the Laplacian operator; x is the distance from the stapes along the BM with $x=0$ at the base (or the stapes); and y is the vertical distance from the BM, with $y=0$ at the BM. By definition, the velocity potential is related to the fluid velocity's components in the x and y directions: $V_x=-\partial\phi/\partial x$ and $V_y=-\partial\phi/\partial y$ [5].

The BM's response to both the pressure difference ($P_{\rm d}$) between the fluid ducts[3] and the OHC forces ($F_{\rm OHC}$) can be described as

$$P_{\rm d}(x)+F_{\rm OHC}(x)=S(x)\delta(x)+\beta(x)\dot{\delta}(x)+M(x)\ddot{\delta}(x) \quad (1)$$

where $S(x)$, $\beta(x)$, and $M(x)$ are, respectively, the BM's stiffness, damping, and mass (per unit area) and δ is the BM's downward displacement. The pressure difference is given by $P_{\rm d}=\rho\,\partial(\phi_{\rm SV}(x,y,t)-\phi_{\rm ST}(x,y,t))/\partial t=2\rho\dot{\phi}$, evaluated at the BM ($y=0$); ρ is the fluid density.

The $F_{\rm OHC}$ term combines forward and backward OHC forces [Fig. 1(b)], described as in [8]

$$F_{\rm OHC}(x)=\alpha S(x)\{\gamma\mathcal{T}(\delta(x-d))-\mathcal{T}(\delta(x+d))\} \quad (2)$$

where α represents OHC motility, expressed as a fraction of BM stiffness, and γ is the ratio of forward to backward coupling, representing relative strengths of OHC forces exerted on the BM segment directly through a Deiters' cell (DC) on which the OHC sits (first term), and indirectly via a phalangeal process (PhP) attached to the reticular lamina (RL) (second term). $\delta(x-d)$ and

[3]Only the scala vestibuli and the scala tympani are considered, since Reissner's membrane, which separates the scala vestibuli and the scala media, is extremely thin, presenting negligible acoustic impedance [23].

$\delta(x+d)$ are the displacements of adjacent upstream and downstream BM segments ($i-1$ and $i+1$), respectively; d denotes the tilt distance, the horizontal displacement between the source and the recipient of the OHC force, assumed to be equal for the forward and backward cases. The function $\mathcal{T}$ models saturation of OHC forces, a nonlinearity evident in physiological measurements [23].

The forward and backward coupling forces' opposite signs account for the fact that OHCs move the BM and the RL in opposite directions. Forward coupling, proposed by others [24], [25], posits that the OHC's basal tilt results in Segment $i-1$'s BM motion reinforcing that of Segment i. Backward coupling, the novel component of ABC, posits that the PhP's apical tilt results in Segment $i+1$'s motion opposing that of Segment i. Adding ABC to a passive model makes the peak in BM displacement higher and sharper, similar to the difference between a live and dead cochlea. This frequency-selective amplification arises because ABC makes the damping negative when the wavelength becomes short (see [8] and [17] for further details).

IV. Circuit Design

Based on the mathematical cochlear model, we design a 2-D nonlinear active cochlear circuit in analog VLSI, taking advantage of the 2-D nature of Si chips. We start by synthesizing a passive model, and then extend it to a nonlinear active one by including ABC with saturation.

A. Passive Cochlear Circuit

The model consists of two fundamental parts: 1) the cochlear fluid and 2) the BM. First, we design the fluid circuit by using the discrete version of Laplace's Equation (in 1-D for simplicity)

$$\frac{1}{\Delta x}\left(\frac{\phi_{n+1}-\phi_n}{\Delta x}-\frac{\phi_n-\phi_{n-1}}{\Delta x}\right)=0$$

where $\phi_n = \phi(n\Delta x)$. The velocity potential may be represented in one of two ways: If node n's voltage represents ϕ_n (voltage-mode), resistors connect adjacent nodes. If node n's voltage represents the $\log(\phi_n)$ (log domain), subthreshold MOS transistors (diffusors) connect adjacent nodes [26], [27]. The latter is simpler to implement (see Fig. 7): We used nMOS transistors for the diffusors and pMOS transistors to take the antilog, yielding a current I_ϕ that is proportional to ϕ_n^κ, a good approximation if the pMOS transistors' κ (subthreshold slope-coefficient) is close to one.

Second, we design a BM segment and, thus, the BM. If I_ϕ represents $2\rho\phi$ (the velocity potential scaled by the fluid density) and I_{mem} represents $\dot{\delta}$ (BM velocity), the BM boundary condition (1) can be expressed as

$$\dot{I}_\phi = S(x)\int I_{\mathrm{mem}}\mathrm{dt}+\beta(x)I_{\mathrm{mem}}+M(x)\dot{I}_{\mathrm{mem}}.$$

(The F_{OHC} term is dealt with in Section IV-C.) Taking the first time derivative and working in the s-domain ($s=j\omega$) yields

$$I_\phi s^2 = S(x)I_{\mathrm{mem}}+\beta(x)I_{\mathrm{mem}}s+M(x)I_{\mathrm{mem}}s^2. \quad (3)$$

We synthesized this second-order system from two low-pass filters (LPFs) by using a custom *Mathematica* program to find a decomposition that provides economy and accommodates ABC readily (see Section IV-C)

$$\begin{aligned}\tau_1 I_{\mathrm{s}}s+I_{\mathrm{s}} &= -I_\phi+I_{\mathrm{o}}\\ \tau_2 I_{\mathrm{o}}s+I_{\mathrm{o}} &= I_\phi-bI_{\mathrm{s}}\\ I_{\mathrm{mem}} &= I_\phi+I_{\mathrm{s}}-I_{\mathrm{o}}.\end{aligned} \quad (4)$$

The LPFs' outputs are I_{s} and I_{o} (a.k.a., state variables); their time constants are τ_1 and τ_2, respectively; b is a gain factor.

The BM's velocity matches the fluid's; thus, we must ensure that

$$I_{\mathrm{mem}} = -\frac{I_\phi(x,\Delta y)-I_\phi(x,0)}{2\rho\Delta y}$$

(recall that, by definition, $V_y = -\partial\phi/\partial y$). The RHS is proportional to the current in the diffusor that connects these two nodes. Therefore, we can satisfy this constraint simply by connecting the BM circuits' current output (I_{mem}) to the fluid circuit—and setting the diffusors' gate voltage (V_{fld}; see Fig. 7) appropriately (i.e., $\rho \propto e^{-\kappa V_{\mathrm{fld}}/\mathrm{u_T}}$, where κ is the nMOS transistors' subthreshold slope coefficient and $\mathrm{u_T}$ is the thermal voltage, 25.6 mV at room temperature).

B. Circuit Analogs of Biology

Given (4), I_ϕ, I_{s}, and I_{o} can be expressed in terms of the output current I_{mem}

$$\begin{aligned}I_\phi &= \frac{(b+1)+(\tau_1+\tau_2)s+(\tau_1\tau_2)s^2}{\tau_1\tau_2 s^2}I_{\mathrm{mem}}\\ I_{\mathrm{s}} &= -\frac{1}{\tau_1 s}I_{\mathrm{mem}}\\ I_{\mathrm{o}} &= \frac{(b+1)+\tau_1 s}{\tau_1\tau_2 s^2}I_{\mathrm{mem}}.\end{aligned} \quad (5)$$

By comparing the expression for I_ϕ with the design target (3), we obtain the circuit counterparts

$$S(x)=\frac{b+1}{\tau_1\tau_2},\ \beta(x)=\frac{\tau_1+\tau_2}{\tau_1\tau_2},\ \text{and}\ M(x)=1 \quad (6)$$

where the mass is normalized. These analogies require that the time constants (τ_1 and τ_2) increase exponentially to simulate the exponentially decreasing BM stiffness (and damping). b allows us to achieve a larger quality factor (a measure of frequency selectivity) for a given choice of τ_1 and τ_2 (limited by capacitor size C: $\tau = C\mathrm{u_T}/\kappa I_\tau$, where I_τ is the current level). That is

$$Q=\frac{\sqrt{S(x)M(x)}}{\beta(x)}=\frac{\sqrt{b+1}}{\sqrt{\frac{\tau_1}{\tau_2}}+\sqrt{\frac{\tau_2}{\tau_1}}}.$$

These circuit-biology relationships help determine the parameter values used in circuit simulation and chip operation.

C. Adding Active Bidirectional Coupling

We synthesize an active BM segment by following the same procedure we used for the passive one, but with the F_{OHC} term included. The design target equation becomes

$$\begin{aligned}I_\phi s^2 &= S(x)I_{\mathrm{mem}}+\beta(x)I_{\mathrm{mem}}s+M(x)I_{\mathrm{mem}}s^2\\ &-\alpha S(x)\left\{\gamma\mathcal{T}\left(\frac{I_{\mathrm{mem}}(x-d)}{s}\right)s-\mathcal{T}\left(\frac{I_{\mathrm{mem}}(x+d)}{s}\right)s\right\}.\end{aligned}$$

We find $I_{\rm mem}/s$ (i.e., the time-integral) by observing that the state variable $I_{\rm s}$ in the passive design (5) is related to $I_{\rm mem}$ by $I_{\rm s} = -I_{\rm mem}/(\tau_1 s)$. Thus

$$\frac{I_{\rm mem}(x-d)}{s} = -\tau_{\rm 1f} I_{\rm sf}, \quad \frac{I_{\rm mem}(x+d)}{s} = -\tau_{\rm 1b} I_{\rm sb}$$

where $I_{\rm sf}$ and $I_{\rm sb}$ represent the output currents, and $\tau_{\rm 1f}$ and $\tau_{\rm 1b}$ are the time constants of the first LPF in the upstream and downstream BM segment, respectively. We replace $\tau_{\rm 1f}$ and $\tau_{\rm 1b}$ by τ_1—the receiving segment's time constant—a good approximation due to the small change in τ between neighboring segments. Therefore, the design target becomes

$$I_{\rm in}s^2 = S(x)I_{\rm mem} + \beta(x)I_{\rm mem}s + M(x)I_{\rm mem}s^2 - r_{\rm ff}S_0(x)\mathcal{T}(-I_{\rm sf})s + r_{\rm fb}S_0(x)\mathcal{T}(-I_{\rm sb})s \quad (7)$$

where $S_0(x) = S(x)\tau_1 = (b+1)/\tau_2$ [see (6)]—τ_1 was factored out by rescaling $\mathcal{T}$; $r_{\rm ff} = \alpha\gamma$ and $r_{\rm fb} = \alpha$ denote the forward and backward OHC force factors, respectively.

We synthesized the circuit following a procedure similar to that used in the passive design. Only the second equation changed

$$\tau_2 I_{\rm o}s + I_{\rm o} = I_\phi - bI_{\rm s} - a_{\rm ff}\mathcal{T}(-I_{\rm sf}) + a_{\rm fb}\mathcal{T}(-I_{\rm sb}) \quad (8)$$

where $a_{\rm ff} = r_{\rm ff}(b+1)$ and $a_{\rm fb} = r_{\rm fb}(b+1)$. Note that to include ABC, we need to only add two currents to the input of the second LPF in each BM segment circuit; these currents are from its adjacent neighbors. Specifically, $I_{\rm sf}$ and $I_{\rm sb}$ are the output currents ($I_{\rm s}$) of the first LPF in the upstream (basal) and downstream (apical) BM segments, respectively.

V. Circuit Implementation

Based on our synthesized design, we implement a Class AB log-domain circuit for the BM segment. We employ the log-domain filtering technique [28] to realize current-mode operation. In addition, we adopt Class AB operation to increase dynamic range, reduce the effect of mismatch, and lower power consumption ([29]–[31]). This differential signaling is inspired by the biological cochlea—the BM's displacement is driven by the pressure difference across it. We present the transistor-level schematics in this section as well as an Si neuron that converts the segment's output current into a pulse stream.

A. Basilar Membrane Circuit

Taking a bottom-up approach, we start by designing a Class AB LPF, a building block for the BM circuit. An LPF is described by

$$\tau I_{\rm out}s + I_{\rm out} = I_\phi$$

where I_ϕ is the input current, $I_{\rm out}$ is the output current, and τ sets the time constant. Its differential counterpart is

$$\tau(I^+_{\rm out} - I^-_{\rm out})s + (I^+_{\rm out} - I^-_{\rm out}) = I^+_\phi - I^-_\phi$$

where each signal is expressed as the difference between its positive ($+$) and negative ($-$) components. The common-mode constraint is

$$\tau I^+_{\rm out}I^-_{\rm out}s + I^+_{\rm out}I^-_{\rm out} = I^2_{\rm q}$$

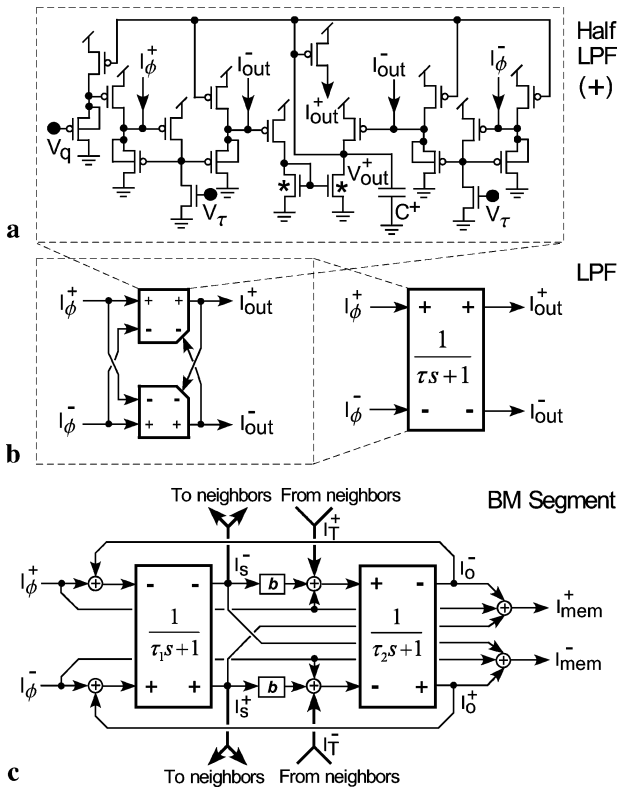


Fig. 2. Low-pass filter (LPF) and basilar membrane (BM) segment circuits. (a) Half-lowpass—filter circuit (* denotes complex current mirror). (b) Complete LPF circuit formed by two half-LPF circuits. (c) BM-segment circuit. It consists of two LPFs and connects to its neighbors, sending current $I_{\rm s}$, and receiving current $I_{\rm T}$ (last two terms in (8), corresponding to ABC). In (b) and (c), current splitting implies current copying.

where $I_{\rm q}$ sets the geometric mean of the output current's components.

Combining the common-mode constraint with the differential design equation yields the positive path's nodal equation (the negative path has superscripts $+$ and $-$ swapped) [29]

$$C\dot{V}^+_{\rm out} = I_\tau \frac{\left((I^+_\phi - I^-_\phi) + \left(\frac{I^2_{\rm q}}{I^+_{\rm out}} - I^+_{\rm out}\right)\right)}{(I^+_{\rm out} + I^-_{\rm out})}.$$

This nodal equation suggests the half-LPF circuit shown in Fig. 2(a). $V^+_{\rm out}$, the voltage on the positive capacitor (C^+), gates a pMOS transistor to produce the corresponding current signal $I^+_{\rm out}$ ($V^-_{\rm out}$ and $I^-_{\rm out}$ are similarly related). The bias $V_{\rm q}$ sets the quiescent current $I_{\rm q}$ while V_τ determines the current I_τ, which is related to the time constant by $\tau = C{\rm u_T}/\kappa I_\tau$. Two of these subcircuits, connected in push-pull, form a complete LPF [Fig. 2(b)]. Specifically, when the input I^+_ϕ charges C^+, it also discharges C^-; similarly, I^-_ϕ charges C^- and discharges C^+.

The BM-segment circuit [Fig. 2(c)] is implemented by using two LPFs interacting in accordance with the synthesized design equations. $I_{\rm mem}$ is the sum of three signals I_ϕ, $I_{\rm s}$, and $I_{\rm o}$ (4). The positive and negative components of $I_{\rm s}$, $I^+_{\rm s}$, and $I^-_{\rm s}$ are the differential output currents of the first LPF (with time-constant τ_1), corresponding to $I^+_{\rm out}$ and $I^-_{\rm out}$ in the LPF symbol [see Fig. 2(b)], respectively. Similarly, $I^+_{\rm o}$ and $I^-_{\rm o}$ are the output currents of the second LPF (with time-constant τ_2).

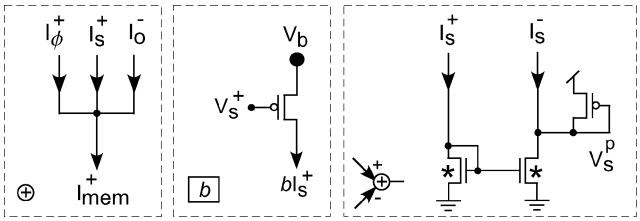


Fig. 3. Summing, scaling, and subtracting circuits. Left: Summing circuit. Tying three wires together sums their currents. Middle: Scaling circuit. Setting $V_{\rm b} = V_{\rm dd} + {\rm u_T}\ln(b)/\kappa$ scales the current ($I_{\rm s}^{+}$) made by a pMOS transistor with the same gate voltage ($V_{\rm s}^{+}$) by b. Right: Subtracting circuit. Tying two wires to either side of a complex current mirror (*) subtracts their currents; passing the result through a diode-connected transistor rectifies it. Superscript p corresponds to $I_{\rm s}^{+} - I_{\rm s}^{-}$, while m corresponds to $I_{\rm s}^{-} - I_{\rm s}^{+}$ (see Fig. 4).

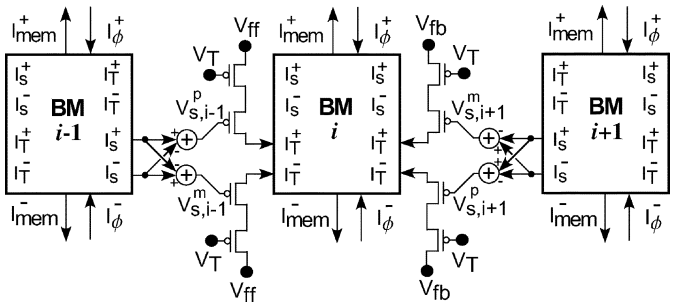


Fig. 4. Active-coupling circuitry. BM Segment i receives a scaled and saturated version of Segment $i-1$ and $i+1$'s $I_{\rm s}$ outputs. The circuitry that couples Segment i to Segments $i-1$ and $i+1$ is omitted for clarity. $V_{\rm ff}$ and $V_{\rm fb}$ set the gain of feedforward and feedbackward coupling, respectively; $V_{\rm T}$ sets their saturation levels.

Summing (+) is implemented by exploiting Kirchhoff's Current Law (Fig. 3); scaling (b) is implemented by biasing a pMOS transistor's source voltage (Fig. 3).[4]

ABC is implemented by exchanging currents between neighboring BM segments (Fig. 4). Each BM sends out $I_{\rm s}$ and receives $I_{\rm T}$, a saturated and scaled version of its neighbor's $I_{\rm s}$ (8). The saturation is accomplished by a current-limiting transistor, which yields $I_{\rm T} = \mathcal{T}(I_{\rm s}) = I_{\rm s}I_{\rm sat}/(I_{\rm s} + I_{\rm sat})$[32], where $I_{\rm sat}$ is set by a bias voltage $V_{\rm T}$. We used a subtract circuit (Fig. 3) to take the difference first because saturation is applied to the differential signal, not to its positive and negative components. The scaling corresponds to the gain factors $a_{\rm ff}$ and $a_{\rm fb}$ in (8), implemented by biasing a pMOS transistor's source voltage ($V_{\rm ff}$ and $V_{\rm fb}$ in Fig. 4(b), respectively).

B. Spiral Ganglion Cell Circuit

In the biological cochlea, the BM's velocity, sensed by inner hair cells (IHCs), is encoded by spiral ganglion cells (SGCs). Behaving like pulse-frequency modulators, SGCs convey information about sound stimuli—including frequency, level, and timing—over the AN (axons of SGCs). Their spikes are evoked by neurotransmitter released from IHCs, each of which drives 10 to 30 SGCs, increasing from apex to base [33]. In our Si cochlea, this fanout is 6.

The IHC circuit has three functional components (Fig. 5): 1) a current mirror takes the difference between $I_{\rm mem}$'s positive and negative components; 2) a current splitter half-wave rectifies the difference, and 3) a class A log-domain LPF filters the half-wave-rectified currents. The bias voltages ($V_{\rm gain}$ and $V_{\rm sat}$) can be varied to yield distinct rate-level relations (i.e., sound level to spike rate).

[4]In the case of $I_{\rm mem}$, the sum is mirrored twice to produce additional copies to feed to the scanner and pulse-frequency modulators.

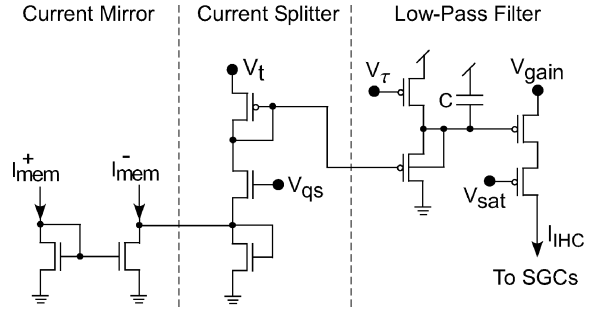


Fig. 5. Inner hair cell circuit. It converts differential currents, representing BM velocity, into a low-pass filtered, half-wave rectified current that drives the spiral ganglion cell (SGC) circuits. $V_{\rm qs}$ sets the quiescent level, $V_{\rm t}$ sets the current splitter's gain, V_{τ} sets the low-pass filter's time-constant, $V_{\rm gain}$ sets the synaptic efficacy, and $V_{\rm sat}$ sets the maximum amplitude.

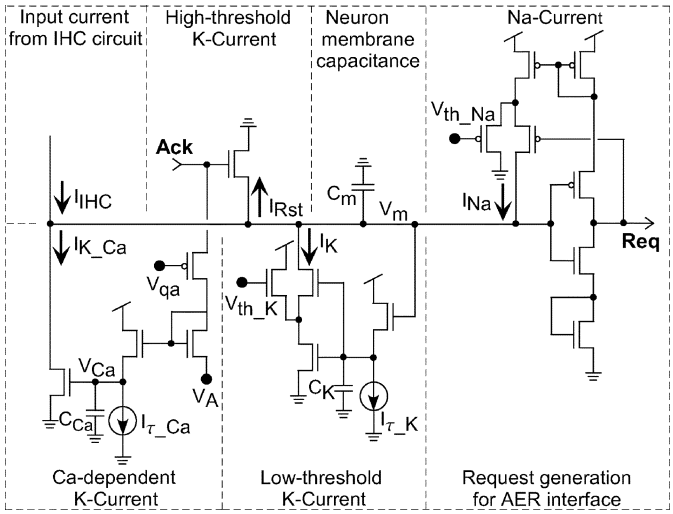


Fig. 6. Spiral ganglion cell circuit. It models three membrane-voltage-dependent currents ($I_{\rm Na}$, $I_{\rm Rst}$, and $I_{\rm K}$) and one Ca-concentration-dependent current ($I_{\rm K\_Ca}$). These currents turn on when V_m exceeds certain levels (set by $V_{\rm th\_Na}$ and $V_{\rm th\_K}$) or when a spike occurs (activates Req, which activates Ack). $I_{\tau\_{\rm K}}$ sets $I_{\rm K}$'s time-constant; $V_{\rm qa}$ and $I_{\tau\_{\rm Ca}}$ set the Ca-concentration's increment (per spike) and time constant, respectively.

Augmenting its static rate-level relation, an SGC's dynamic properties enhance the encoding of a sound stimulus' temporal features: It fires at a higher rate at stimulus onset, due to the presence of a Ca-concentration-dependent K-current [34]. And, from cycle to cycle, it is more likely to fire when the sinusoid is rising most rapidly (phase locking [35]), due to the presence of a low-threshold K-current [36], [37]. In addition to these two K-currents, the SGC circuit models an Na current that generates an all-or-none spike and a high-threshold K-current that resets the membrane [38] (Fig. 6).

An address-event encoder transmits the SGC circuits' spikes off-chip [39]–[41]. To communicate with the address-event encoder, the SGC makes a request when it spikes and clears this signal when acknowledged (see Req and Ack in Fig. 6). The spike is encoded as a unique address (specifying row and column). The receiver chip decodes this address and delivers the spike to the target neuron.

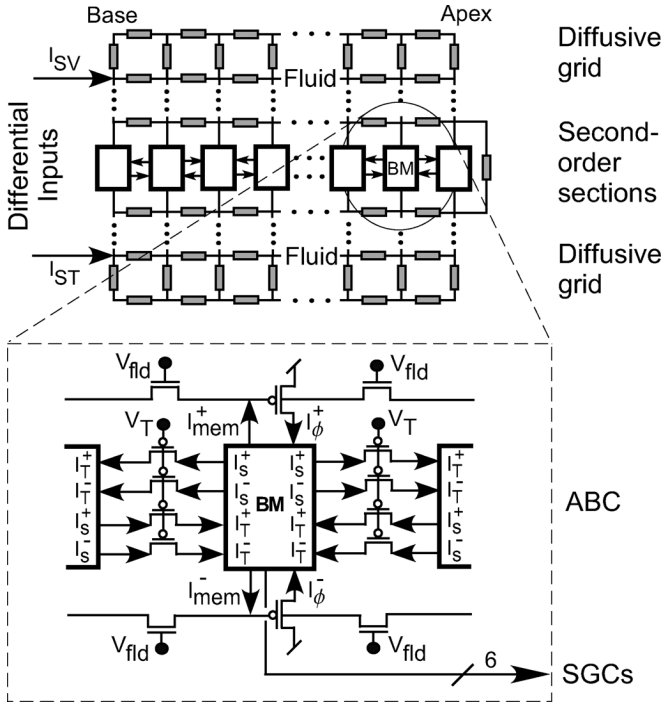


Fig. 7. Active-coupling architecture. Differential audio signals are applied at the base of two diffusive-element grids (representing the top and bottom fluid, respectively; $V_{\rm fld}$ determines the fluids' density), connected at the apex by a single diffusive element (representing the helicotrema). Second-order sections (representing the BM segments) embedded between the grids send/receive currents to/from their immediate neighbors ($V_{\rm T}$ as in Fig. 4), realizing active bidirectional coupling (ABC). Their output currents ($I_{\rm mem}$, representing BM velocity) each drive six pulse-frequency-modulation circuits (representing the spiral ganglion cells (SGCs)).

C. Chip Architecture

We fabricated a version of this design with 360 BM-segment circuits, two 4680-element (360×13) fluid grids, and 2160 (360×6) SGC circuits (Fig. 7). The number of BM segments was chosen to satisfy the requirements of our software simulation—the chip has approximately 55 segments per octave (assuming a 200–20 kHz range, or 6.6 octaves). The fluid-element grids' height (13) was chosen to match the biological cochlea's aspect ratio, a factor important in controlling the traveling wave's behavior [5]. The number of SGC circuits per BM segment (6) was chosen to ensure that the stimulus evokes multiple spikes per cycle—an octave-wide response will produce up to 33 kspikes/s (assuming a maximum spike frequency of 100 Hz). A die photo of the chip is shown in Fig. 8.

VI. Chip Responses

We measured the Si cochlea's BM responses to pure tones and its AN responses to complex sounds. To supply sinusoidal current as input, we applied the logarithm of a half-wave-rectified sinusoid to the top and bottom fluid grids; these two voltage signals were 180° out-of-phase. Scaled to match the pMOS transistors' subthreshold slope-coefficient ($\kappa = 0.58$, measured), the half-wave rectified voltage signals' peak amplitude varied from 0.12 to 0.36 V, dropping from a baseline of 2.26 V ($V_{\rm dd} = 2.4$ V), in 0.04-V steps. These values correspond to input current amplitudes of 0 to 48 dB, increasing in 8-dB steps, with 24 dB corresponding to a medium sound level. We set up the cochlea chip's time-constant-setting voltages ($V_{\tau 1}$ and $V_{\tau 2}$) by tuning the base and the apex to approximately 20 kHz and slightly below 200 Hz, respectively. Linear interpolation (implemented with two polysilicon lines spanning the Si cochlea's length) gave rise to exponentially decreasing time-constant currents. The saturation level of ABC currents was set to its maximum level ($V_{\rm T} < 1.7$ V, putting the pMOS transistor above threshold) unless otherwise stated.

We measured frequency responses as well as longitudinal responses. To obtain frequency responses, we swept the input frequency and measured BM current outputs (from both positive and negative paths) at a particular segment.[5] To obtain longitudinal responses, we kept the input frequency fixed and measured current outputs at consecutive segments along the cochlea's length. Selecting a particular segment or sweeping through consecutive ones is realized with a built-in scanner (modified from [42] to accommodate snaking). AN responses, on the other hand, were measured in parallel by capturing (time-stamped) address-events over a universal-serial-bus (USB) link.

Here, we present frequency responses measured from linearly spaced BM segments and longitudinal responses to octave-spaced pure tones, both at an input level of 24 dB. We also map the dependence of frequency and signal-to-noise ratio (SNR) on position, also at a 24-dB input level. In addition, we present frequency responses obtained at various input intensities (0 to 48 dB), demonstrating automatic gain control. We then demonstrate the role of ABC by disabling it. Finally, we present the chip's real-time responses to a chirp-click sound sequence.

A. Frequency Responses

Frequency responses reveal the tuning of individual BM segments (Fig. 9).[6] Despite some irregularities in response shape and peak height (due to transistor mismatch), the chip's responses captured the characteristics of the biological responses, at least qualitatively. Frequency responses are peaked and cutoff slope is steep (more so in some segments than others), with peak or characteristic frequencies (CFs) ranging from 13.8 k to 218 Hz for these six BM segments (40-segment spacing). Phase accumulates gradually at first, then more rapidly near the peak [marked by dots in Fig. 9(b)], and plateaus after the peak. The large accumulation indicates a traveling wave; the plateau indicates its extinction.

Histograms of measurements from 12 equally spaced BM segments reveal marked differences across the cochlea, indicating poor parameter matching among segments, to the extent that desired performance was not achieved at all taps (Fig. 10). Tip-to-tail ratios (amplitude difference between the peak and lowest frequency point), a commonly used measure of cochlear amplification, ranges from 9 to 45 dB, approaching the chinchilla's performance (53 dB). CF phase ranges from 0.4 to -3.5π radians, spanning the chinchilla's performance (-1.67π radians). Q_{10} ranges from 0.1 to 2.7, reaching the chinchilla's performance (2.55 at medium sound levels). The cutoff slope ranges from -2 to -54 dB/octave, falling 1.6 times or more short of the chinchilla (-85 dB/octave). In addition,

[5]Segment numbers increase from base to apex, starting from 1.

[6]Segments beyond 240 were not considered because they did not respond robustly, probably due to the large discontinuity we observed at each U-turn (Segments 60, 120, 180, etc.), presumably caused by doping-level deviations at the array's edges (dummy cells were not deployed).

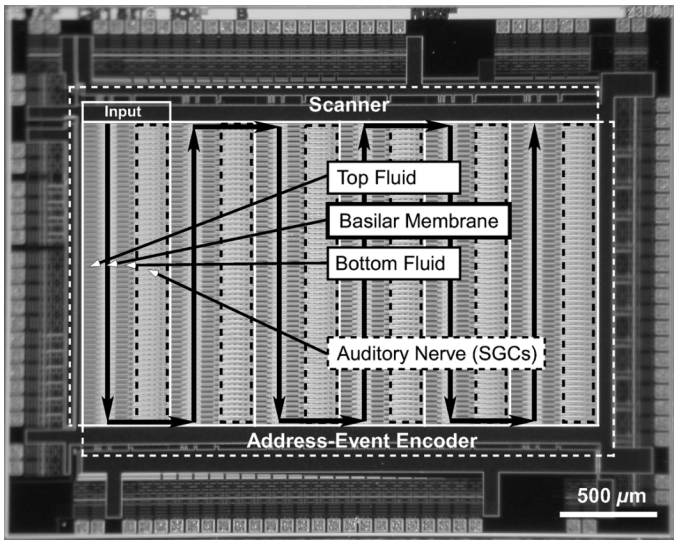


Fig. 8. Die photo. Fabricated in a 5M 1P 0.25-μm CMOS process, the ABC cochlea, with six 60-segment columns snaking to yield a desirable aspect ratio, occupies 10.9 mm^2. Input, basilar membrane, top/bottom fluid, auditory nerve (axon of spiral ganglion cells, or SGCs), scanner, and address-event encoder circuits are labeled.

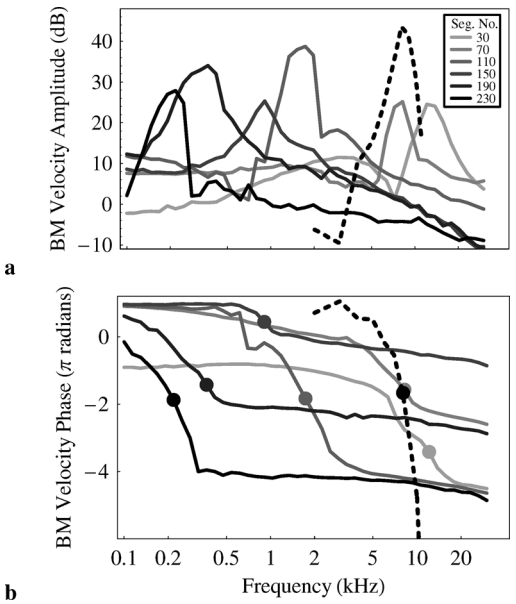


Fig. 9. Frequency responses of six BM segments, spaced 40 segments apart, from 30 to 230 (24-dB input level). (a) Amplitude. (b) Phase (dots mark the characteristic frequencies). Biological data are provided for comparison (dashed line, chinchilla measurement at the medium sound level [43]).

similar to the chinchilla cochlea's basal region, the Si cochlea has a logarithmic frequency-position map: Segment number (n) is related to CF (f, in Hertz) by $\log_{10} f = 4.26 - 0.0082n$ [Fig. 11(a)].

To evaluate noise accumulation in our ABC architecture, we calculated the SNR at each of the 12 equally spaced cochlear

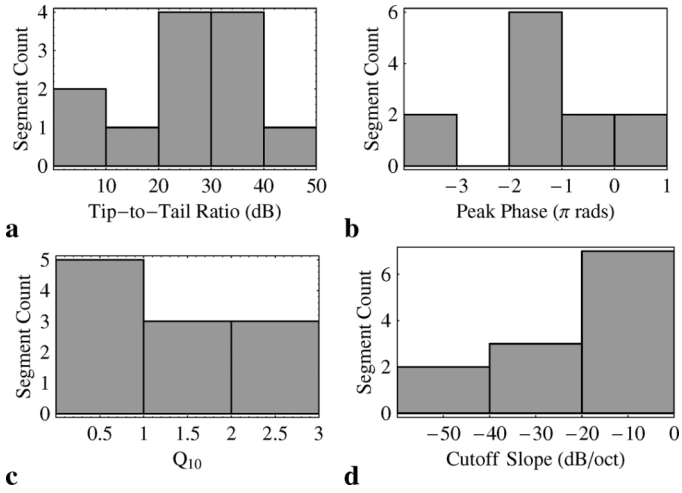


Fig. 10. Histograms of measurements from 12 BM segments, spaced 20 segments apart, from 10 to 230 (24-dB input level). (a) Tip-to-tail ratio. (b) Peak phase. (c) Q_{10}. (d) Cutoff slope.

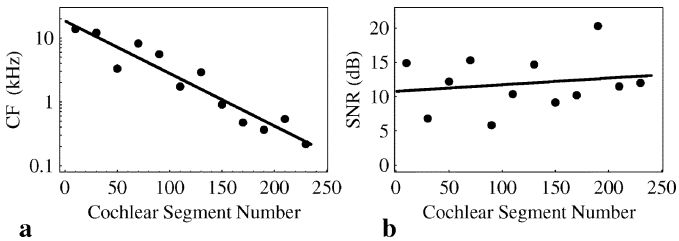


Fig. 11. Frequency-position map and signal-to-noise ratios (SNRs) at the 24-dB input level. (a) The frequency a segment responds maximally to (CF; dots) is logarithmically related to its position (line). (b) SNR at 12 cochlear segments (for CF). Dots: Data; line: Linear regression.

segments, stimulated at their CF [Fig. 11(b)]. SNR was computed as the ratio between the signal's power (i.e., squared amplitude at the CF) and the noise's power (sum of squared amplitude at all frequencies—see Fig. 14(a)). A linear regression of SNR versus cochlear position n yielded $\mathrm{SNR(dB)} = 0.01n + 10.7$, indicating insignificant accumulation.

B. Longitudinal Responses

Longitudinal responses give a snapshot of the entire basilar membrane, thereby providing a direct measurement of the traveling wave, whose wavelength ABC is sensitive to. They also show how the wave's amplitude builds up as it travels from the base to the apex, providing evidence that ABC acts in a distributed fashion. The chip's longitudinal responses show large variations from segment to segment (due to mismatch), which we filtered with a 10-segment moving average in order to estimate the response characteristics [Fig. 12(a)].

We measured longitudinal responses to four pure tones, with octave spacing [Fig. 12(b)]. A 4-kHz tone elicits a peak response at Segment 85 (characteristic place, or CP) while a 500-Hz tone travels further and peaks at Segment 178. The CPs for the two intermediate frequencies (1 and 2 kHz) are Segment 166 and 139, respectively. Tip-to-tail ratios range from 12 to 32 dB; Q_{10}s

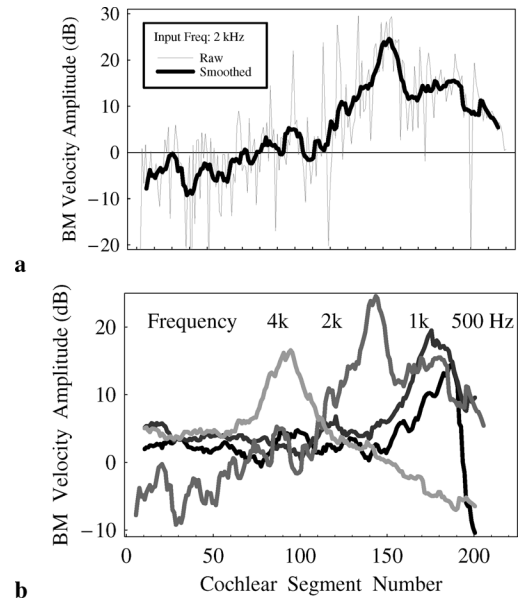


Fig. 12. Longitudinal responses. (a) Raw and smoothed longitudinal responses (2 kHz tone input at 24 dB). A 10-segment moving average removes the large segment-to-segment variations. (b) Smoothed longitudinal responses to four octave-space frequencies. As frequency increases (from 500 to 4 k Hz), the response peaks closer to the Si cochlea's base.

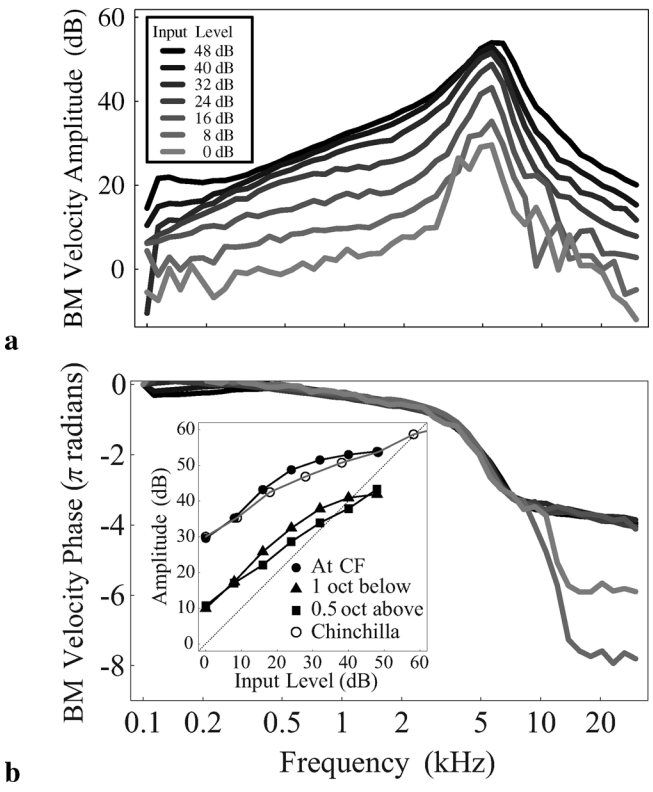


Fig. 13. Nonlinear compression. BM-velocity frequency responses for different input amplitudes (Segment 100; $\mathrm{CF} = 5.6$ kHz). (a) Amplitude. Equally spaced responses indicate linear behavior. (b) Phase. Inset: Input–output functions, measured at CF, an octave below, and half an octave above. Biological measurement is provided for comparison (open circles, chinchilla measurement from [43], shifted to align the lowest input level tested with that of the chip). Dotted line: Identity ($y = x$).

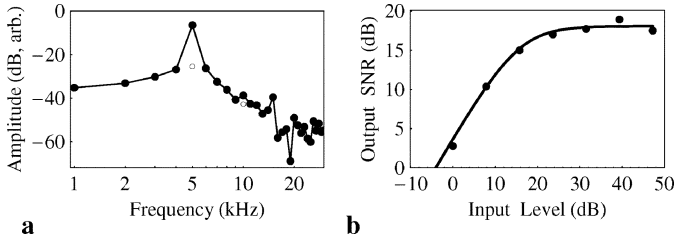


Fig. 14. Signal-to-noise ratio (Segment 100). (a) Spectra of output to near-CF input (5-kHz tone input at 40 dB). Interpolation (unfilled dots) was used to estimate noise at the input frequency and its harmonics. (b) Output SNR increased with increasing input amplitude at low intensities but saturated above 24 dB. *Fit*: Michaelis–Menten function.

range from 0.9 to 1.2; and cutoff slopes range from -16 to -70 dB/octave.[7]

C. Input/Output Functions

Input/output (I/O) functions reveal the Si cochlea's nonlinear behavior. We increased the input amplitude exponentially by increasing the voltage applied linearly, calculating the corresponding amplitude (in decibels) based on the chip's subthreshold slope-coefficient (measured experimentally). The amplitude range applied was constrained on the high side by strong inversion (leaving the subthreshold region), and on the low side by the noise floor. We set $V_{\mathrm{T}} = 1.9$ V to saturate active coupling at the upper end of this range, thereby producing compression.

BM responses show compressive growth, first at the CF and then at nearby frequencies (Fig. 13). As a result, BM responses become more broadly tuned with increasing input amplitude; Q_{10} drops from 1.8 to 1.1. There is a corresponding decrease in cutoff slope, which drops from -44 to -13 dB/octave. Unlike biology, where there is a basal shift (to lower frequency) [44], the CF hardly changes, probably due to insufficiently high input levels. Response phase does not change significantly; this is the case in biology as well. The larger phase plateaus (exactly 2π apart) at low input amplitudes (0 and 8 dB) are due to noisy responses in the cutoff region.

Compression does not occur symmetrically around the peak: It sets in at lower intensities for frequencies below the CF (see Fig. 13, inset). Whereas at the CF (5.6 kHz), compression sets in when the input amplitude exceeds 24 dB, one octave below (2.8 kHz) it occurs at 32 dB, and half an octave above (7.9 kHz), it occurs at 48 dB (the largest amplitude applied). This result suggests that upstream segments (higher CFs) contribute to automatic gain control, more so than downstream segments (lower CFs).

The chip's CF behavior agrees qualitatively with the chinchilla measurements (see Fig. 13, inset), except that at high intensities, which the chip input did not reach, the chinchilla's I/O function became linear again, resembling a passive cochlea. Above or below the CF, the chip's I/O functions are less linear (more compressive) than the chinchilla's (data not shown), presumably because the chip's tuning is broader so that compression at high sound levels occurs with a larger spread.

To find the lowest detectable input amplitude, we measured SNR at the output (defined as the signal-squared over noise-

[7]The number of segments spanned by an octave was calculated from the CF range of the first 240 segments.

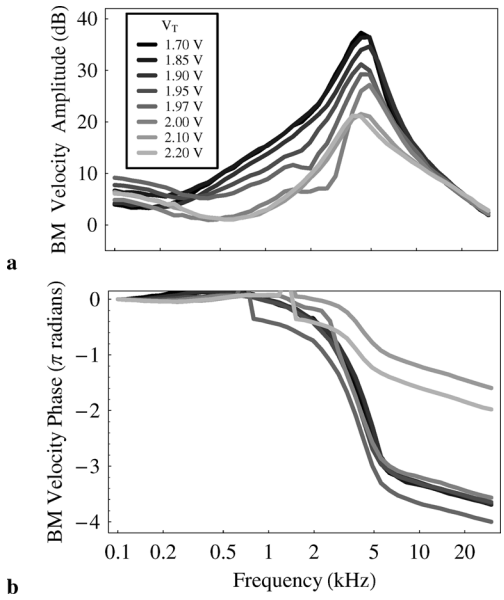


Fig. 15. Frequency responses for various coupling saturation levels (Segment 100; 16-dB input level). (a) Amplitude. (b) Phase. The saturation level decreases as V_{T} increases.

squared) for various input amplitudes and extrapolated to 0 dB (Fig. 14). Output SNR increased from 2.8 to 19 dB as input amplitude increased from 0 to 48 dB. A Michaelis–Menten function fitted the SNR's initial increase and asymptotic behavior well: $R = R_m x^n/(x^n + {R_h}^n)$, with $R_m = 64.3$, $R_h = 5.9$, and $n = 1.8$, where R and x represent the SNR and input amplitude (relative to the smallest current applied), respectively.[8] Extrapolating the fit yields an output SNR of 1 (i.e., 0 dB) at an input amplitude of -4 dB, indicating a 52-dB input dynamic range.

D. Effect of Active Bidirectional Coupling

Varying the coupling's saturation level (through V_{T}) demonstrates the ABC's role. In all responses presented thus far, except for Section VI-C, the saturation level was high enough to avoid saturation. It gets progressively lower as V_{T} (which gates pMOS transistors) increases, producing saturation at lower and lower input levels. Coupling is negligible for $V_{\mathrm{T}} = 2.2$ V, which corresponds to a passive cochlea.

We obtained a series of frequency responses from Segment 100 with different saturation levels (Fig. 15). Decreasing saturation levels resulted in smaller response amplitudes. The amplitude decreased monotonically from 33.3 to 15.1 dB (arbitrary scale) at the CF, an 18 dB drop. Since decreases were more prominent in this region, responses became more broadly tuned; Q_{10} decreased monotonically from 1.14 to 0.45. The phase did not change significantly—except for the weakest couplings.

We also measured longitudinal responses to a 2-kHz tone with ($V_{\mathrm{T}} = 0.26$ V) and without ($V_{\mathrm{T}} = 2.35$ V) coupling (Fig. 16). The peak amplitude was 14.6 dB larger with coupling; the cutoff slope was 35 dB/octave steeper; Q_{10} increased from 0.39 to 1.16. These increases are comparable to those seen in Segment 100's frequency response (increases of 18.2 dB, 22.0 dB/octave, and 0.69, respectively).

[8]To convert to decibels, take $\log_{10}$ of x or R and multiply by 20 or 10, respectively.

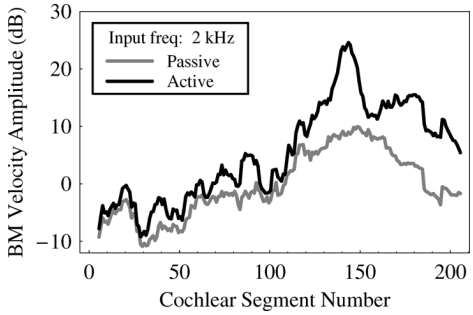


Fig. 16. Longitudinal responses with (active; black) and without (passive; gray) coupling (2-kHz tone input at 24 dB).

In summary, ABC increases gain and sharpens tuning, achieving responses that are qualitatively comparable to physiological measurements. Indeed, the cases with and without ABC resemble live (active) and dead (passive) cochlea, respectively; thus, ABC captures the role of OHC electromotility—at least qualitatively.

E. Si Auditory Nerve

We visualized the Si AN's response by constructing a cochleagram (Fig. 17). This raster plot displays spike trains of all SGCs in Segments 1 to 240, a total of 1440 (240×6) outputs,[9] with time running from left to right and the segment number running from top (base) to bottom (apex)—high to low frequency.

The Si AN responds to the chirp-click sequence with a wave of spike activity followed by a flash (see Fig. 17). The wave propagates from the base to the apex in response to the chirp's decreasing frequency. It becomes more sharply defined after the first 60 segments (i.e., 360 SGCs), indicating the extent to which frequency selectivity arises cooperatively. The flash lights up all outputs simultaneously in response to the click's broad frequency content, except for apex, where it is masked by the chirp's close proximity in time (contiguity). This masking is due to SGC spike-rate adaptation, which emphasizes sound onsets. A few highly excitable SGCs (e.g., Channel 223 and 480) respond throughout most of the stimulus; this behavior is due to transistor mismatch.

In summary, the Si AN encodes a sound's frequency, intensity, and timing. It uses a place code for frequency: only neurons at a certain location fire. It uses a rate code for intensity: these neurons spike at higher rates. And it uses a real-time code for timing: spike rates change in real time, with sound onset emphasized by SGC spike-rate adaptation. The chip's specifications are summarized in Table I.[10]

VII. Discussion

The chip measurements presented here demonstrate that ABC overcomes the major shortcomings of previous Si cochlea architectures, summarized in Table II. In the cascade architecture, noise increased a hundredfold (asymptoting after 30 segments) [3]. With ABC, noise does not accumulate, as demonstrated by our SNR measurements [Fig. 11(b)]. In the passively coupled

[9]SGCs from Segment 241 to 360 (apical third of the cochlea) were omitted for the same reason stated earlier.

[10]Mean $\pm$ standard deviation is quoted for $n = 12$ measurements.

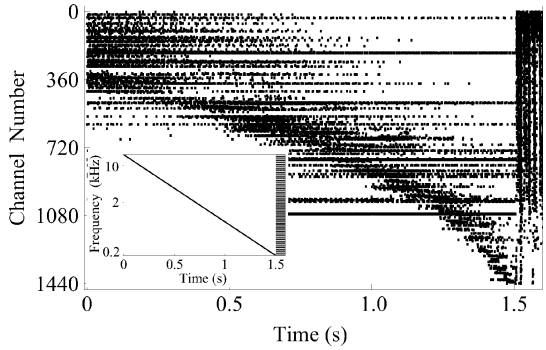


Fig. 17. Chirp-click cochleagram. The chirp invokes a wave that propagates from the base (top) toward the apex (bottom). The click invokes a flash that lights up all but the lowest frequency outputs. Inset: Chirp-click sound. On a logarithmic scale, the 1.5-s chirp's frequency decreases linearly and the 0.1-s click's 50 discrete frequencies are equally spaced. Both span 16 kHz to 200 Hz.

TABLE I
Si COCHLEA SPECIFICATIONS

Specification	*Values*
Fabrication process	5M 1P 0.25 μm CMOS
Chip die dimensions	3.76 mm $\times$ 2.91 mm
Transistors	115,000
Cochlear segments	360
SGC spike outputs	2,160
Frequency range	200 $\sim$ 20k Hz
Input dynamic range	52 dB
Signal-to-noise ratio	11.9$\pm$4.0 dB
Tip-to-tail ratio	26.9$\pm$11.7 dB
Phase at peak	-1.38 ± 1.31 π radians
Q_{10}	1.16$\pm$0.92
Cut-off slope	-19.6 ± 17.5 dB/octave
Compression at peak	24 dB
Supply voltage	A$V_{\rm dd}$ = 2.4 V; D$V_{\rm dd}$ = 2.5 V
Total power	51.8 mW (35.9 mW for analog core)

TABLE II
Si COCHLEA ARCHITECTURES

Architecture	*Major issues*
Cascade [2], [12], [45]	Noise and delay accumulate; poor fault tolerance.
Passive coupling [5], [6], [46]	Destructive interference occurs due to large phase changes at resonance.
Active coupling [this work]	Reflections occur when device mismatch causes abrupt changes.

architecture, gain decreased by 25 dB [6]. With ABC, this destructive interference is overcome, as demonstrated by our gain and tuning measurements [Figs. 15 and 16]. However, ABC's gain increase was limited to 18 dB by mismatch-induced traveling-wave reflections. We confirmed that these reflections can reduce gain and broaden tuning by performing simulations with mismatch included.

The dominant source of transistor mismatch is threshold-voltage variation, which has been shown to be Gaussian distributed, with variance inversely proportional to the transistor's channel area [47]. When transistors operate in weak inversion for low power consumption, their currents are log normally distributed. For instance, in a 0.35-μm CMOS process, the currents' coefficient of variation (standard deviation over mean), or CV, is 9.2% and 22% for 11.4 $\times$ 11.4 λ and 4.6 $\times$ 4.6 λ ($\lambda = 0.18$ μm) nMOS transistors, respectively [48]. Given the transistors sizes in our circuits (Table III), we used log normally

TABLE III
Si COCHLEA DEVICE SIZES

Circuit	*Transistor W/L* ($\lambda = 0.12$ μm)	*Circuit*	*Capacitor Size (fF)*
BM Mirror	20/20	BM	4,200 (Individual)
Rest of BM	10/10	IHC-LPF	209
Fluid Diffusor	10/10	SGC	100
ABC	10/10		
SGC	8/8; 8/4; 8/2		

distributed parameter values with CVs ranging up to 25% in our simulations (see the Appendix).

We quantified the mismatch's effect on active amplification (tip-to-tail ratio) and tuning sharpness (Q_{10}), extracted from smoothed BM velocity responses (Fig. 18). With OHC motility factor $\alpha = 0.15$, these metrics dropped from 85$\pm$4 (mean $\pm$ standard deviation) to 41$\pm$7 dB and from 5.2 $\pm$ 0.6 to 1.1 $\pm$ 0.8, respectively, as CV increased from 5% to 25%, becoming similar to the passive case ($\alpha = 0$), albeit with substantially larger variance. This loss of sensitivity and selectivity could be counteracted by increasing α. For CV = 25%, increasing α from 0.15 to 0.25 increased peak gain from 41 $\pm$ 7 to 70 $\pm$ 13 dB and Q_{10} from 1.1 $\pm$ 0.8 to 3.6 $\pm$ 1.5, with the variance increasing dramatically in both cases. For comparison, with $\alpha = 0.15$, these metrics were 89 dB and 5.6, respectively, in the absence of mismatch.

These simulation results suggest that mismatch accounts for shortfalls in the chip's overall performance as well as variability among its segments. For the value of α the chip used (0.14—estimated from bias voltages that determine $a_{\rm ff}$, $a_{\rm fb}$, and b), the simulations reproduce the range of values we measured for tip-to-tail ratio and Q_{10} (see Table I) with a parameter CV of slightly above 20% [see Fig. 18(c) and (d)], which is twice the 10% current CV of the chip's (mostly) 10 $\times$ 10 λ transistors.[11] We could not confirm the predicted performance improvement with values of α greater than 0.14 because they produced instability in the chip.

These simulation results also suggest that ABC has the potential to exceed the best performance achieved by Si cochleae to date. Sarpeshkar *et al.*'s hybrid parallel-cascade architecture achieved a tip-to-tail ratio of 77 dB [3]. Whereas, Fragniére's passively coupled architecture achieved a Q_{10} of 2.34 [6]. In comparison, our simulations predict that with the practical OHC motility factor of 0.14, ABC can achieve a tip-to-tail ratio of 85 $\pm$ 4 dB and a Q_{10} of 5.2 $\pm$ 0.6 if parameter-CV is reduced to 10%. This requires decreasing the current CV from 10% to 5% by increasing transistor sizes from 10 $\times$ 10 to 20 $\times$ 20 λ. These larger transistors would only increase the chip's size from 10.9 to 14.3 mm^2, since the BM segments currently occupy only 10.3% of its area (excluding the capacitors). However, this promised performance remains to be proven in Si.

VIII. CONCLUSION

We presented a mixed-signal VLSI implementation of a 2-D nonlinear cochlear model that utilizes a novel cochlear amplifier mechanism, ABC. ABC produces large amplification and sharp tuning to soft sound and nonlinear compression and broad

[11]This doubling could be produced by cascading two current mirrors. The fact that this is less than the actual number of mirroring operations in a BM segment is explained by the averaging that occurs when several segments' outputs are actively and passively coupled together to yield the measured response.

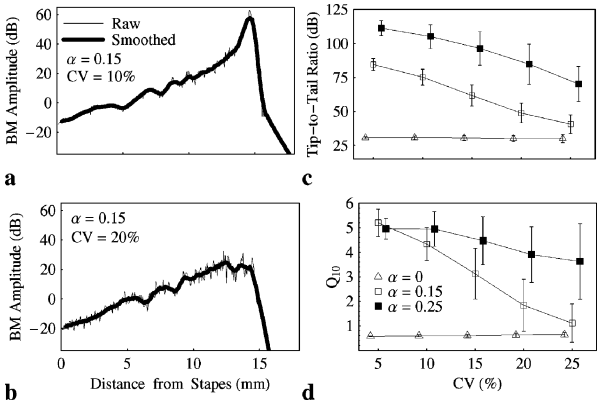


Fig. 18. Simulated responses (2-kHz tone). Parameter coefficient-of-variation (CV) and OHC motility factor (α) were varied. (a) and (b) BM velocity's longitudinal response with $\alpha = 0.15$ and $\mathrm{CV} = 10\%$ and 20%. A ten-segment moving average yielded smoothed responses. (c) and (d) Tip-to-tail ratio and Q_{10} decrease with CV. Datapoints for $\alpha = 0$ and $\alpha = 0.25$ were shifted slightly to avoid overlap with those for $\alpha = 0.15$; error bars give ± 1 standard deviation for 100 trials.

tuning to loud sound. Rather than detecting signal amplitude and implementing an AQC loop, ABC simply mimics OHC-force saturation, which acts instantaneously; and ABC senses the wavelength, which adds frequency selectivity. It successfully combats destructive interference in the coupled architecture.

ABC was vulnerable to reflections that occur when BM properties change abruptly due to device mismatch. However, reducing mismatch by increasing transistor area will decrease the number of BM segments per octave, which will broaden frequency tuning. Alternatively, as mismatch perturbs BM properties at the finest spatial scale, it could be counteracted by increasing the wavelength, something the three-to-five-OHC tilt observed in the Organ of Corti [21], [22] would achieve. The caveat is that lengthening the wavelength will also broaden tuning. Further study is required to determine which approach offers the most favorable tradeoff.

APPENDIX

Here, we describe our simulation procedures. We formulated a linear version (no saturation) of the model described in Section III using circuit variables (i.e., I_ϕ, I_s, I_o, and I_mem) and parameters (i.e., τ_1 and τ_2), based on (4) and (8). We discretized the model horizontally and vertically into a 360×13 grid (identical to the chip). We obtained solutions for the fluid's velocity potential (I_ϕ) and the BM's velocity (I_mem) at each location using the finite difference method [49] in the frequency domain. On each trial, we introduced log normally distributed variations to τ_1 and τ_2. We assumed they were identically distributed for simplicity (i.e., equal mean and variance). The mean was determined based on the biophysical analogs (6). We also applied log normally distributed variation to the cochlear fluid's density (ρ).

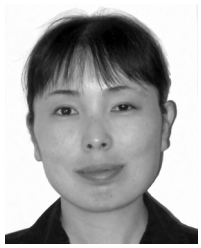
ACKNOWLEDGMENT

The authors would like to thank J. V. Arthur for his help with debugging the chip-PC USB interface used to visualize spike activity. They would also like to thank the Associate Editor T. Delbrück and three anonymous reviewers for their valuable comments and suggestions on improving this presentation.

REFERENCES

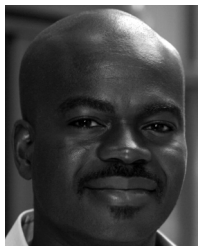
[1] J.-J. Sit and R. Sarpeshkar, "A cochlear-implant processor for encoding music and lowering stimulation power," *IEEE Pervasive Comput.*, vol. 1, no. 7, pp. 40–48, Jan.–Mar. 2008.
[2] R. F. Lyon and C. A. Mead, "An analog electronic cochlea," *IEEE Trans. Acoust. Speech Signal Process.*, vol. 36, no. 7, pp. 1119–1134, Jul. 1988.
[3] R. Sarpeshkar, R. F. Lyon, and C. Mead, "A low-power wide-dynamic-range analog VLSI cochlea," *J. Analog Integr. Circuits Signal Process.*, vol. 16, no. 3, pp. 245–274, 1998.
[4] R. Sarpeshkar, C. Salthouse, J.-J. 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 Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 711–727, Apr. 2005.
[5] L. Watts, "Cochlear mechanics: Analysis and analog VLSI," Ph.D. dissertation, Calif. Inst. Technol., Pasadena, CA, 1993.
[6] E. Fragniére, "A 100-channel analog CMOS auditory filter bank for speech recognition," in *Proc. IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2005, pp. 140–141.
[7] W. E. Brownell, C. R. Bader, D. Bertrand, and Y. Deribaupierre, "Evoked mechanical responses of isolated cochlear outer hair cells," *Science*, vol. 227, no. 4683, pp. 194–196, 1985.
[8] B. Wen and K. A. Boahen, "A linear cochlear model with active bi-directional coupling," in *Proc. 25th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2003, pp. 2013–2016.
[9] B. Wen and K. A. Boahen, "A 360-channel speech preprocessor that emulates the cochlear amplifier," in *Proc. IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, 2006, pp. 556–557.
[10] S. Martignoli, J.-J. van der Vyver, A. Kern, Y. Uwate, and R. Stoop, "Analog electronic cochlea with mammalian hearing characteristics," *Appl. Phys. Lett.*, vol. 91, no. 6, pp. 064108–064108, 2007.
[11] T. Hamilton, C. Jin, J. Tapson, and A. van Schaik, "A 2-D cochlea with Hopf oscillators," in *Proc. IEEE Biomedical Circuits and Systems Conf.*, 2007, pp. 91–94.
[12] R. Sarpeshkar, R. F. Lyon, and C. Mead, "An analog VLSI cochlea with new transconductance amplifiers and nonlinear gain control," in *Proc. IEEE Int. Symp. Circuits and Systems*, Atlanta, GA, 1996, vol. 3, pp. 292–295.
[13] T. J. Hamilton, C. Jin, A. van Schaik, and J. Tapson, "An active 2-D silicon cochlea," *IEEE Trans. Biomed. Circuits Syst.*, vol. 2, no. 1, pp. 30–43, Mar. 2008.
[14] B. C. J. Moore, *An Introduction to the Psychology of Hearing*. New York: Elsevier Academic Press, 2004.
[15] W. S. Rhode and P. H. Smith, "Characteristics of tone-pip response patterns in relationship to spontaneous rate in cat auditory-nerve fibers," *Hearing Res.*, vol. 18, no. 2, pp. 159–168, 1985.
[16] B. Wen and K. A. Boahen, "Active bidirectional coupling in a cochlear chip," in *Advances in Neural Information Processing Systems*, Y. Weiss, B. Schölkopf, and J. Platt, Eds. Cambridge, MA: MIT Press, 2006, vol. 18, pp. 1497–1504.
[17] B. Wen, "Modeling the nonlinear active cochlea: Mathematics and analog VLSI," Ph.D. dissertation, Univ. Pennsylvania, Philadelphia, 2006.
[18] C. D. Summerfield and R. F. Lyon, "ASIC implementation of the lyon cochlea model," in *Proc. IEEE Int. Conf. Acoust. Speech and Signal Process.*, San Francisco, CA, 1992, pp. 673–676.
[19] M. P. Leong, C. T. Jin, and P. H. W. Leong, "An FPGA-based electronic cochlea," *EURASIP J. Appl. Signal Process.*, vol. 7, pp. 629–638, 2003.
[20] C. A. Mead, *Analog VLSI and Neural Systems*. Reading, MA: Addison-Wesley, 1989.
[21] Y. Raphael, M. Lenoir, R. Wroblewski, and R. Pujol, "The sensory epithelium and its innervation in the mole rat cochlea," *J. Comp. Neurol.*, vol. 314, pp. 367–382, 1991.
[22] I. Russell and K. Nilsen, "The location of the cochlear amplifier: Spatial representation of a single tone on the guinea pig basilar membrane," in *Proc. Nat. Acad. Sci.*, 1997, vol. 94, pp. 2660–2664.
[23] C. D. Geisler, *From Sound to Synapse: Physiology of the Mammalian Ear*. Oxford, U.K.: Oxford University Press, 1998.
[24] C. R. Steele, G. Baker, J. Tolomeo, and D. Zetes, "Electro-mechanical models of the outer hair cell," in *Proc. Int. Symp. Biophysics of Hair Cell Sensory Systems*, H. Duifhuis, J. W. Horst, P. van Dijk, and S. M. van Netten, Eds., Singapore, 1993, pp. 207–214.
[25] C. D. Geisler and C. Sang, "A cochlear model using feed-forward outer-hair-cell forces," *Hear. Res.*, vol. 86, pp. 132–146, 1995.
[26] K. A. Boahen and A. G. Andreou, "A contrast sensitive silicon retina with reciprocal synapses," in *Advances in Neural Information Processing Systems*, J. E. Moody and R. P. Lippmann, Eds. San Mateo, CA: Morgan Kaufmann, 1992, vol. 4, pp. 764–772.
[27] A. G. Andreou and K. A. Boahen, "Translinear circuits in subthreshold MOS," *J. Analog Integr. Circuits Signal Process.*, vol. 9, no. 2, pp. 141–166, 1996.

[28] D. R. Frey, "Log-domain filtering: An approach to current-mode filtering," *Proc. Inst. Elect. Eng., G Circuits Devices Syst.*, vol. 140, no. 6, pp. 406–416, 1993.
[29] K. Zaghloul and K. A. Boahen, "An on-off log-domain circuit that recreates adaptive filtering in the retina," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 52, no. 1, pp. 99–107, Jan. 2005.
[30] T. Y. W. Choi, B. E. Shi, and K. A. Boahen, "An on-off orientation selective address event representation image transceiver chip," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 2, pp. 342–353, Feb. 2004.
[31] W. K. Chen, *VLSI Handbook*. Boca Raton, FL: CRC, 2000.
[32] T. Delbrück, "Bump circuits for computing similarity and dissimilarity of analog voltages," in *Proc. Int. Joint Conf. Neural Networks*, Seattle, WA, 1991, vol. 1, pp. I–475–479.
[33] M. C. Liberman, L. W. Dodds, and S. Pierce, "Afferent and efferent innervation of the cat cochlea: Quantitative analysis with light and electron microscopy," *J. Comp. Neurol.*, vol. 301, no. 3, pp. 443–460, 1990.
[34] C. L. Adamson, M. A. Reid, Z. L. Mo, J. Browne-English, and R. L. Davis, "Firing features and potassium channel content of murine spiral ganglion neurons vary with cochlear location," *J. Comp. Neurol.*, vol. 447, no. 4, pp. 331–350, 2002.
[35] J. H. Wittig and K. A. Boahen, "Silicon neurons that phase-lock," in *Proc. IEEE Int. Symp. Circuits and Systems*, 2006, pp. 4535–4538.
[36] J. Santos-Sacchi, "Voltage-dependent ionic conductances of type-I spiral ganglion-cells from the guinea-pig inner-ear," *J. Neurosci.*, vol. 13, no. 8, pp. 3599–3611, 1993.
[37] E. J. Moore, D. B. Hall, and T. Narahashi, "Sodium and potassium currents of type I spiral ganglion cells from rat," *Acta Oto-Laryngologica*, vol. 116, no. 4, pp. 552–560, 1996.
[38] K. Yamaguchi and H. Ohmori, "Voltage-gated and chemically gated ionic channels in the cultured cochlear ganglion neurone of the chick," *J. Physiol.*, vol. 420, pp. 185–206, 1990.
[39] J. Lazzaro, M. Wawrzynek, M. Mahowald, M. Sivilotti, and D. Gillespie, "Silicon auditory processors as computer peripherals," *IEEE Trans. Neural Netw.*,, vol. 4, no. 3, pp. 523–528, May 1993.
[40] K. A. Boahen, "A burst-mode word-serial address-event channel-I: Transmitter design," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 7, pp. 1269–1280, Jul. 2004.
[41] V. Chan, S.-C. Liu, and A. van Schaik, "AER EAR: A matched silicon cochlea pair with address event representation interface," *IEEE Trans. Circuits Syst. I*, vol. 54, no. 1, pp. 48–59, Jan. 2007.
[42] C. A. Mead and T. Delbrück, "Scanners for visualizing activity of analog VLSI circuitry," *J. Analog Integr. Circuits Signal Process.*, vol. 1, pp. 93–106, 1991.
[43] M. A. Ruggero, N. C. Rich, S. S. Narayan, and L. Robles, "Basilar membrane responses to tones at the base of the chinchilla cochlea," *J. Acoust. Soc. Am.*, vol. 101, no. 4, pp. 2151–2163, 1997.
[44] L. Robles and M. A. Ruggero, "Mechanics of the mammalian cochlea," *Physiol. Rev.*, vol. 81, no. 3, pp. 1305–1352, 2001.
[45] A. van Schaik, E. Fragniére, and E. A. Vittoz, "Improved silicon cochlea using compatible lateral bipolar transistors," in *Advances in Neural Information Processing Systems*, D. S. Touretky, M. C. Mozer, and M. E. Hasselmo, Eds. Cambridge, MA: MIT Press, 1996, vol. 8, pp. 671–677.
[46] A. van Schaik and E. Fragniére, "Pseudo-voltage-domain implementation of a design of two-dimensional silicon cochlea," in *Proc. IEEE Int. Symp. Circuits and Systems*, 2001, vol. 3, pp. 185–188.
[47] P. R. Kinget, "Device mismatch and tradeoffs in the design of analog circuits," *IEEE J. Solid-State Circuits*, vol. 40, no. 6, pp. 1212–1224, Jun. 2005.
[48] B. Linares-Barranco, T. Serrano-Gotarredona, R. Serrano-Gotarredona, and G. Vicente-Sanchez, "On mismatch properties of MOS and resistors calibrated ladder structures," in *Proc. IEEE Int. Symp. Circuits and Systems*, 2004, vol. 1, pp. 377–380.
[49] S. T. Neely, "Finite-difference solution of a two-dimensional mathematical model of the cochlea," *J. Acoust. Soc. Am.*, vol. 69, no. 5, pp. 1386–1393, 1981.



Bo Wen (SM'05–M'06) received the B.S. and M.S. degrees in electrical engineering from the University of Science and Technology Beijing, China, in 1995 and 1999, respectively, and the Ph.D. degree in neuromorphic engineering from the University of Pennsylvania, Philadelphia, in 2006.

Her research interest is investigating neural coding/mechanisms of sound and speech signals in the auditory system through computational modeling—including computer software and mixed-signal VLSI circuits—and neurophysiological recordings.



Kwabena Boahen (M'06) received the B.S. and M.S.E. degrees in electrical and computer engineering from the Johns Hopkins University, Baltimore, MD, in 1989, and the Ph.D. degree in computation and neural systems from the California Institute of Technology, Pasadena, in 1997.

Currently, he is an Associate Professor in the Bioengineering Department at Stanford University, Stanford, CA. He is a Bioengineer who uses silicon integrated circuits (ICs) to emulate the way neurons compute, linking the seemingly disparate fields of electronics and computer science with neurobiology and medicine. His contributions to the field of neuromorphic engineering include a silicon retina that could be used to give the blind sight and a self-organizing chip that emulates the way the developing brain wires itself up. His scholarship is widely recognized, with many publications to his name, including a cover story in the May 2005 issue of *Scientific American*.

Dr. Boahen has received several distinguished honors, including a Fellowship from the Packard Foundation in 1999, a CAREER award from the National Science Foundation in 2001, a Young Investigator Award from the Office of Naval Research in 2002, and the National Institutes of Health Director's Pioneer Award in 2006. From 1997 to 2005, he was on the faculty of the University of Pennsylvania, Philadelphia, where he held the first Skirkanich Term Junior Chair.