

# Silicon Neurons That Phase-Lock

John H. Wittig Jr and Kwabena Boahen

Department of Bioengineering, University of Pennsylvania, Philadelphia, Pennsylvania 19104  
 {jwittig, boahen}@seas.upenn.edu

**Abstract**—We present a silicon neuron with a dynamic, active leak that enables precise spike-timing with respect to a time-varying input signal. Our neuron models the mammalian bushy cell, which enhances the phase-locking of its acoustically driven inputs. Our model enhances phase-locking by up to 38% (quantified by vector strength) across a 60 dB range of acoustic intensities, and up to 22% over a passive leak. Its conductance-based log-domain design yields a compact and efficient circuit, fabricated in 0.25  $\mu\text{m}$  CMOS, that is an ideal timing-enhancing component for neuromorphic speech recognition systems.

## I. IMPROVING SPEECH RECOGNITION SYSTEMS

Automatic speech recognition (ASR) has been sought by engineers for more than 50 years [1]. Unlike human speech recognition, which degrades gracefully under progressively deteriorating acoustic conditions, ASR systems are sensitive to variance in the sound source and to background noise. In contrast, human listeners improve recognition in noisy environments by incorporating knowledge of a speaker’s location, or the pitch of his or her voice (i.e., active listening and the cocktail party problem [2]). To do so, they exploit the speech waveform’s fine temporal structure [3], discarded by ASR preprocessors, which instead extract the instantaneous energy (but not phase) of numerous frequency channels across the spectrum [1]. By not only including, but enhancing timing information during preprocessing, ASR systems may access the plethora of sound-cues that humans rely on.

Custom-designed hardware models of the auditory periphery are well suited to perform ASR preprocessing—they are power efficient yet can evaluate biologically-inspired computations in real-time. Several silicon cochleas have been created that frequency-decompose sound by modeling the observed mechanical behavior of the inner ear [4,5]. Furthermore, inspired by the biological auditory periphery, circuit designers have mimicked the analog-to-digital conversion that transpires at the auditory nerve (AN) by coding the cochlear output as a spike rate [6,7].

In addition to using a spike-rate code to represent the spectral energy of sound (as decomposed along the length of the basilar membrane), the mammalian AN represents the phase of an acoustic waveform with spike timing. Sound localization computations that compare phase differences between the two ears require 10  $\mu\text{s}$  timing accuracy, which exceeds the precision of individual AN fibers. Indeed, AN spike timing (as well as other acoustic features) is enhanced in the mammalian cochlear nucleus (CN) [8]. In the cat, Joris et al. recorded from CN cells, and found that, for frequencies less than 1 kHz, globular bushy cells’ spike timing was more precise than their AN inputs, with a sparse code of a single spike per cycle [9].

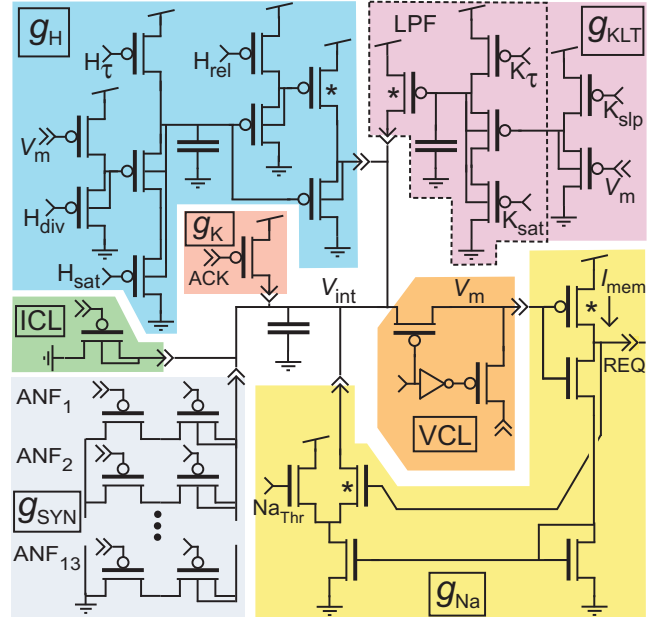


Fig. 1. Bushy cell circuit.  $V_m$ , the membrane voltage, drives  $g_H$ , the hyperpolarization-activated mixed-cation conductance;  $g_{KLT}$ , the low-threshold potassium conductance; and  $g_{Na}$ , the spike-generating sodium conductance. REQ, the spike request, is acknowledged by ACK, which drives  $g_K$ , the spike-resetting potassium conductance.  $ANF_i$ , the auditory nerve input, drives  $g_{SYN}$ , the synaptic conductance. ICL, the current-clamp transistor, is an excitatory input controlled off chip. Currents produced by these conductances are integrated on  $V_{int}$ , which can be isolated from  $V_m$  by VCL, the voltage-clamp circuitry. The low-pass filter building block is outlined (LPF).

The only VLSI model of a CN neuron used a “moderately high” leak and a fast refractory period to realize a bushy cell-like response (spikes aligned at the onset of sound) [10]. With a fixed leak, however, only a narrow range of stimulus intensities elicit this response—below which the neuron fails to respond due to excessive inhibition and above which the neuron spikes repetitively during the on-phase of a stimulus cycle. In contrast, when a biological bushy cell is stimulated with intense sound its active conductances dynamically limit repetitive action-potentials (until the bombardment of synaptic input relaxes during the off-phase of the stimulus cycle). Physiologists identified a large potassium conductance that activates near rest,  $g_{KLT}$  [11], which has been shown to enhance phase-locking [12]. We show here that  $g_{KLT}$  is also responsible for intensity adaptation.

Taking a neuromorphic approach to achieving spike-timing enhancement, we have implemented a silicon model of the

globular bushy cell that includes  $g_{KLT}$ . In Section II, we describe our bushy cell circuit. In Section III, we characterize the circuit with static and time-varying stimuli, then demonstrate that  $g_{KLT}$  enables our bushy cell circuit to phase-lock to an acoustic stimulus over a wide intensity range. In Section IV, we summarize our results and their implications for extracting both sound source pitch and location for ASR.

## II. NEURON IMPLEMENTATION

The heart of our silicon bushy cell is a log-domain low pass filter (LPF) whose output controls four active conductances (Fig. 1). We represent the cell’s membrane voltage as the current flowing through a subthreshold pMOS device operating in saturation ( $I_{mem} \propto e^{\kappa(V_{dd}-V_m)/u_T}$ ). Our simplified implementation of the conductance equation ( $(\frac{C_{mem} u_T/\kappa}{I_{g_{KLT}}+I_{g_H}}) \frac{dI_{mem}}{dt} + I_{mem} \propto \frac{I_{g_{SYN}}+I_{g_{Na}}+I_{g_H}}{I_{g_{KLT}}+I_{g_H}/I_{H_{rel}}}$ ) assumes the relatively fast spike-generating and resetting conductances ( $g_{Na}$  and  $g_K$ ) do not contribute appreciably to the membrane time-constant.

Each active conductance contributes to the cell’s response. The fast-activating sodium conductance ( $g_{Na} \propto I_{mem}^2$ ) is responsible for generating spikes; the moderately fast-activating potassium conductance that has a low activation-threshold ( $g_{KLT} \propto I_{mem}/I_{K_{slp}}$ ) is responsible for enhancing timing; and the slow, hyperpolarization-activated conductance ( $g_H \propto I_{H_{div}}/I_{mem}$ ) is responsible for setting the resting potential while decreasing the membrane time-constant. The proportionality constants (gains) are set by the bias voltages  $Na_{Thr}$ ,  $K_{slp}$ , and  $H_{div}$ , respectively. Additionally, the fast-activating potassium conductance ( $g_K$ ) is included to reset the neuron after it spikes.

These active conductances’ dependence on  $I_{mem}$  is realized using translinear (subthreshold CMOS) circuits, and their dynamics with LPFs, except for  $g_{Na}$ , which responds instantaneously; it simply uses positive feedback circuitry [13].  $g_H$  and  $g_{KLT}$  use variants of the LPF with a translinear input stage that implements the equations above, and limiting transistors that set the maximum conductance. By driving the integration node ( $V_{int}$ ) with current sources, both  $g_H$  (scaled) and  $g_{KLT}$  dynamically control the membrane time-constant. Additionally,  $g_H$  sets the resting level through a source follower.

We designed each conductance’s subcircuit to be flexible yet compact. For  $g_{KLT}$ ’s dependence on  $I_{mem}$ ,  $K_{slp}$  and  $K_{sat}$  adjust the gain and the saturation level, respectively. We well-connected the  $V_m$ -driven transistor to neutralize kappa ( $\approx 0.6$  in this process), which would otherwise cause a square root-like increase in  $g_{KLT}$  at low  $I_{mem}$  levels. For  $g_H$ ’s dependence,  $H_{div}$  effectively shifts the value of  $V_m$  where limiting occurs.  $H_{rel}$  balances  $g_H$ ’s relative contributions to the neuron’s time-constant and resting level. For  $g_{Na}$ ’s dependence,  $Na_{Thr}$  shifts the gain and onset (i.e., spiking threshold). These biases enabled us to readily match the biological g-v curves.

Our silicon bushy cell can be driven by voltage-clamp, current-clamp, or silicon AN synapses. To voltage-clamp the cell, we tie  $V_m$  directly to an off-chip voltage source after

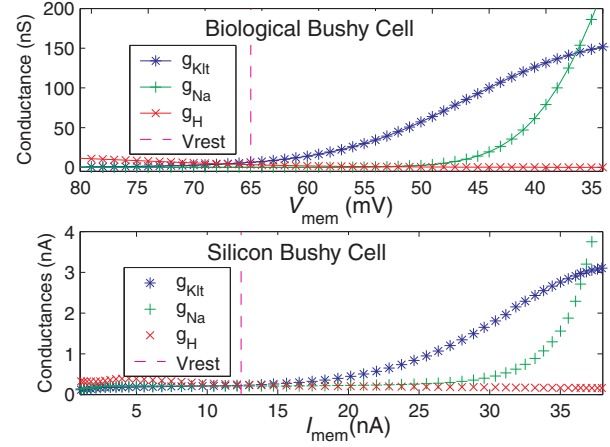


Fig. 2. Modeling bushy cell conductances. Activation curves for three conductances in the biological bushy cell (top) and in our silicon bushy cell (bottom). Bias settings:  $K_{slp} = 2.3$  V;  $K_{sat} = 2.0$  V;  $H_{div} = 1.4$  V;  $H_{sat} = 1.775$  V;  $H_{rel} = 2.1$  V;  $Na_{thr} = 1.0$  V;  $V_{dd} = 2.5$  V.

isolating it from the active conductances’ integration node ( $V_{int}$ ) by turning the voltage-clamp transistor off (VCL). To current-clamp the cell, we apply an off-chip voltage that represents the log of the input current to  $V_{int}$  through a source-follower (ICL). To apply synaptic input, we route 0.1 ms pulses to the thirteen excitatory inputs ( $ANF_i$ ).

Over one thousand copies of our silicon bushy cell were included in a 4410-neuron CN chip that we designed, submitted, and tested. With 544,692 transistors in  $11.33$  mm<sup>2</sup>, it was fabricated in TSMC’s  $0.25$   $\mu$ m CMOS process. In addition to 1080 bushy cells (15% of the cat’s population), the chip includes four other CN cell types that enhance different acoustic features [8]. AN and CN spikes are communicated digitally using the Address Event Representation (AER) [14], and routed to and from a desktop computer via an AER-USB2.0 link. Individual currents (Fig. 1, \*) are measured using a scanner and scaled to compensate for a pad gain of 1000 (estimated).

## III. NEURON CHARACTERIZATION

### A. Conductance-Voltage Curves

Using voltage-clamp, we tuned our silicon bushy cell’s active conductances to qualitatively match those observed biologically. Rothman and Manis [15] formulated g-v curves after characterizing many bushy cells *in vitro* (Fig. 2 top). In steady-state, their model bushy cell rests at  $-66$  mV. At potentials hyperpolarized to rest only  $g_H$  is active. At potentials slightly depolarized from rest,  $g_{KLT}$ , the “low-threshold” conductance, activates and eventually saturates as the spike-generating  $g_{Na}$  becomes active. Our silicon bushy cell’s g-I curves match the biological ones (Fig. 2 bottom); it rests at  $I_{mem} \approx 12$  nA.

We model  $g_{KLT}$  and  $g_H$  with static time-constants, whereas Rothman and Manis use a bell-shaped voltage dependence. In their model,  $g_{KLT}$ ’s time-constant ranges from 1.5 to 6.4 ms, peaking near rest. They did not measure the time-constant of  $g_H$ , but in another CN neuron (octopus cell) it ranges from 30

to 175 ms [16]. Measuring the time-constants of each silicon conductance in voltage-clamp, we set  $g_{\text{KLT}}$ 's to 3 ms ( $K_{\tau} = 2.29$  V) and  $g_{\text{H}}$ 's to 80 ms ( $H_{\tau} = 2.4$  V).

Together,  $g_{\text{KLT}}$  and  $g_{\text{H}}$  dynamically set the neuron's membrane time-constant, yielding a 1 ms time-constant at rest in Rothman and Manis's model. Measuring the membrane time-constant in current-clamp by injecting a small excitatory pulse that does not significantly activate  $g_{\text{KLT}}$ , we set our bushy cell's membrane time-constant to 1.2 ms at rest by scaling  $g_{\text{KLT}}$  and  $g_{\text{H}}$ 's g-I curves.

To probe the role of  $g_{\text{KLT}}$ , we compare results when the conductance is enabled, disabled, and substituted with a (large) passive leak. The disabled setting converts  $g_{\text{KLT}}$  to a fast spike-reset conductance by reducing its input sensitivity ( $K_{\text{slp}} = 2.0$  V) and time-constant ( $K_{\tau} = 2.0$  V) while increasing its maximum output ( $K_{\text{sat}} = 1.0$  V). We simulate a passive leak by increasing  $g_{\text{H}}$  ( $H_{\text{rel}} = 2.5$  V), in addition to disabling  $g_{\text{KLT}}$ .

### B. Response to current injection

Our silicon bushy cell follows time-varying current-clamp input with both  $I_{\text{mem}}$  and its spike timing (Fig. 3 *top*). Similar to the current-clamp response observed experimentally, the membrane responds quickly to a large step input (membrane time-constant of a millisecond), fires a solitary spike, then quickly settles to an elevated level. At the offset of the current step, the membrane rapidly dips below the resting level and then settles back within 20 ms.

The fast response and solitary onset spike are a result of shunting inhibition by  $g_{\text{KLT}}$ . As  $I_{\text{mem}}$  rises,  $g_{\text{KLT}}$  activates and shunts the input current preventing spiking. However, due to a delay, a solitary spike can occur when  $g_{\text{Na}}$  elicits positive-feedback before  $g_{\text{KLT}}$  sufficiently kicks in. We chose a current-clamp level just sufficient to elicit spikes, which occur approximately every other cycle.

When  $g_{\text{KLT}}$  is disabled, neither  $I_{\text{mem}}$  nor spike-times follow the input (Fig. 3 *middle*). As the stimulus turns on and off,  $I_{\text{mem}}$  rises and decays slowly. Over multiple cycles,  $I_{\text{mem}}$  integrates up, eventually reaching spike threshold at a random stimulus phase. We see little relation between stimulus phase and spike-times with  $g_{\text{KLT}}$  disabled.

When a passive leak replaces  $g_{\text{KLT}}$ ,  $I_{\text{mem}}$  again follows the input, though spike-times are not relegated to the stimulus onset (Fig. 3 *bottom*).  $I_{\text{mem}}$  rapidly settles to a level dictated by the difference between leak and input currents, unless the input surpasses the leak, which causes  $I_{\text{mem}}$  to integrate to spike threshold. As the leak does not increase with increasing  $I_{\text{mem}}$ , it cannot limit multiple-spikes during the on-phase of the stimulus. In this case we chose an input level slightly above threshold, which elicits rapid and repetitive spiking throughout the on-phase.

### C. Response to synaptic input

With  $g_{\text{KLT}}$  enabled, our silicon bushy cell accurately follows the phase of a pure-tone acoustic stimulus (Fig. 4 *left*). To test our model's response to sound, we generated stochastic AN

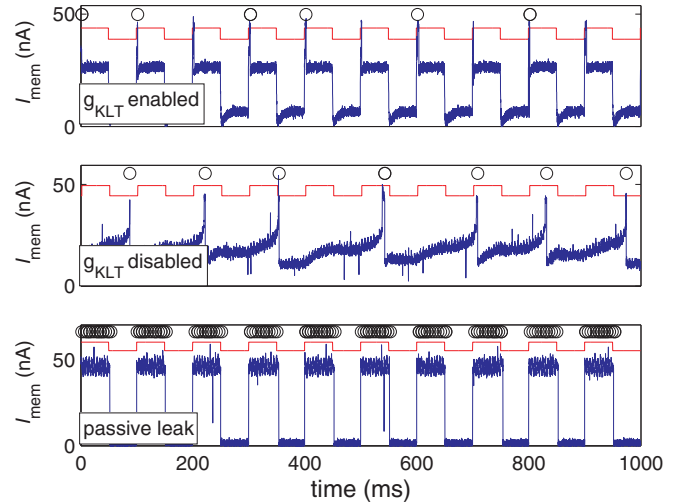


Fig. 3.  $g_{\text{KLT}}$  enhances temporal response to injected current. *top* With  $g_{\text{KLT}}$  enabled,  $I_{\text{mem}}$  follows the 10 Hz square wave's on (370 nA) and off (0 nA) transitions; spikes (o) are relegated to the on transitions. *middle* With  $g_{\text{KLT}}$  disabled,  $I_{\text{mem}}$  is slow to react, has an elevated resting level, and integrates charge over successive current pulses; spikes occur with no apparent phase relation to the stimulus waveform. In this case, we reduced the square wave's amplitude to 28 nA to obtain a similar spike-rate. *bottom* With a passive leak,  $I_{\text{mem}}$  follows the input, has a resting level of zero, and spikes only during the on-phase. At this input level (1.69  $\mu\text{A}$ ), slightly above threshold, spikes occur throughout the on-phase.

spike trains based on a computational model of the guinea-pig cochlea-AN complex [17]. Our bushy cell is innervated by thirteen AN fibers that exhibit a range of spontaneous rates, from 0.5 to 150 Hz. We applied 100 presentations of AN spikes, generated in response to a 25 ms, 250 Hz tone, at 70 dB SPL (with 25 ms between presentations).

Compared to its aggregate AN inputs, the bushy cell fires far fewer spikes overall, has a higher proportion of spikes at sound onset, and fires more precisely in phase with the acoustic stimulus. These characteristics are in line with what we observed when exciting the bushy cell with a square-wave current input:  $g_{\text{KLT}}$  limits spikes to the onset of a strong stimulus and prevents multiple spikes during stimulation (decreasing the overall spike rate).

We quantify phase-locking ability by the magnitude of the normalized vector sum of the spike phases (vector strength, VS) [9]. A VS of one corresponds to perfect phase-locking, where the neuron fires at exactly the same phase on every cycle. A VS of zero corresponds to no phase-locking, where the neuron fires at random phases. The bushy cell exhibits better phase locking (VS = 0.95) than its AN inputs (aggregate VS = 0.82) at 70 dB SPL. In contrast, the bushy cell's VS decreases to 0.90 with a passive leak.

The benefits of  $g_{\text{KLT}}$  (over a passive leak) are most apparent at high intensity levels (Fig. 4. *right*). At 110 dB SPL,  $g_{\text{KLT}}$  makes a dramatic difference: It limits multiple spikes and increases phase-locking (VS = 0.97) despite a decrease in AN phase-locking (VS = 0.60). The passive-leak neuron, on the other hand, fired multiple times per cycle and decreased its phase-locking (VS = 0.75). The superior phase-locking

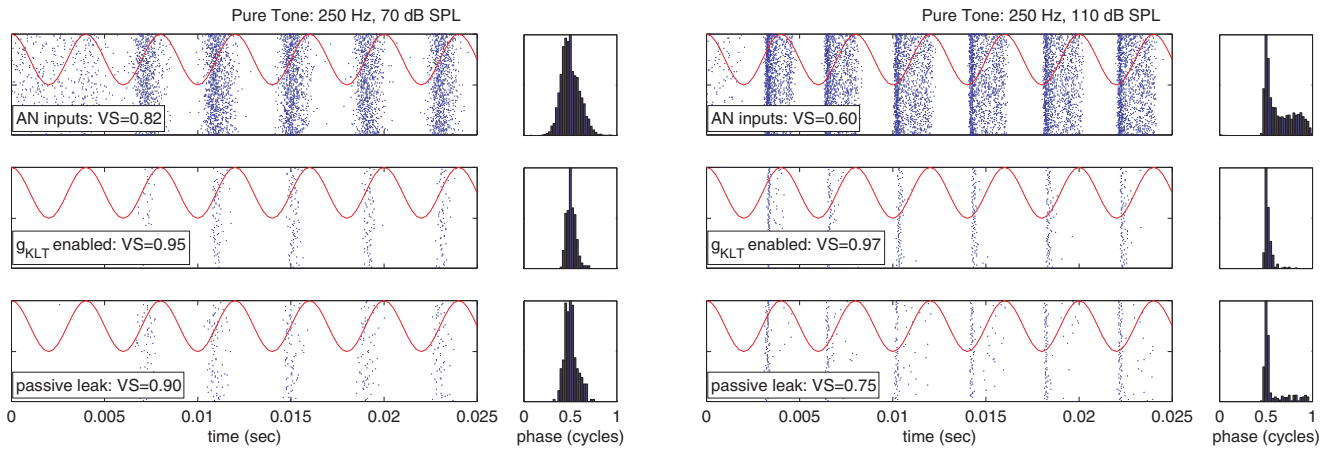


Fig. 4.  $g_{KLT}$  enhances temporal response to synaptic inputs. Spike rasters (100 trials) and phase histograms in response to 250 Hz pure tone (red) at 70 and 110 dB SPL (left and right panels, respectively): *top* simulated AN fibers (all thirteen are collapsed into a single spike train in each trial); *middle* silicon bushy cell with  $g_{KLT}$ ; *bottom* silicon bushy cell with a passive leak. Phase histograms are computed from spikes occurring after 10 ms, and are normalized and shifted to be maximal at 0.5 cycles.

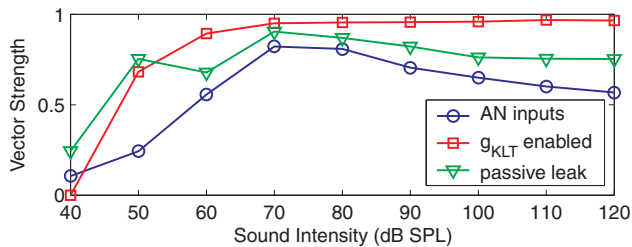


Fig. 5.  $g_{KLT}$  enhances phase-locking over a wide range of intensities. Vector strength values for the aggregate AN, our silicon bushy cell with  $g_{KLT}$ , and with  $g_{KLT}$  replaced by a passive leak.

achieved with  $g_{KLT}$  occurs from 60 to 120 dB SPL (Fig. 5).

#### IV. CONCLUSION

We have implemented a silicon model of a bushy cell, including four active conductances, whose dependence on membrane voltage qualitatively matches those of the biological bushy cell.  $g_{KLT}$  enables the bushy cell to enhance timing over a range of 60 dB SPL. With a passive leak, timing enhancement degrades as intensity increases. Our  $g_{KLT}$  circuit is efficient and compact: It operates with sub-threshold currents using a log-domain low pass filter comprised of only six transistors. By endowing silicon neurons in an auditory preprocessor with this conductance, they will provide precise timing cues for extracting sound source pitch and location, increasing the robustness of ASR systems in real-world acoustic environments.

#### ACKNOWLEDGMENT

We thank Donata Oertel and Paul Manis for sharing their expertise in cochlear nucleus physiology. This work was supported by the MOSIS Educational Program (chip fabrication), the University of Pennsylvania's NIDCD Training Grant in the Neurobiology of Otorhinolaryngology (fellowship to JW), and the Packard Foundation (fellowship to KB).

#### REFERENCES

- [1] N. Morgan, H. Boullard and H. Hermansky, "Automatic Speech Recognition: An Auditory Perspective," in *Speech Processing in the Auditory System*, S. Greenberg, et al., Eds. New York: Springer, 2004, pp.309-338
- [2] S. Haykin and Z. Chen, "The cocktail party problem," *Neural Computation*, vol. 17, pp. 1875-902, 2005
- [3] Z. M. Smith, B. Delgutte and A. J. Oxenham, "Chimaeric sounds reveal dichotomies in auditory perception," *Nature*, vol. 416, pp. 87-90, 2002
- [4] R. F. Lyon and C. Mead, "An analog electronic cochlea," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 36, pp. 1119-1134, 1988
- [5] L. Watts, "Cochlear Mechanics: Analysis and Analog VLSI," California Institute of Technology, PhD Thesis, 1993
- [6] A. McEwan and A. van Schaik, "An alternative analog VLSI implementation of the Meddis inner hair cell model," *IEEE International Symposium on Circuits and Systems*, vol. 4, pp. IV-928-31, 2004
- [7] A. van Schaik and S.-C. Liu, "AER EAR: A Matched Silicon Cochlea Pair with Address Event Representation Interface," *IEEE International Symposium on Circuits and Systems*, pp. 4213-4216, 2005
- [8] E. D. Young and D. Oertel, "Cochlear Nucleus," in *The synaptic organization of the brain*, G. M. Sheperd, Eds. Oxford: University Press, 2004, pp.125-164
- [9] P. X. Joris, et al., "Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency," *J Neurophysiol*, vol. 71, pp. 1022-36, 1994
- [10] A. van Schaik, E. Fragniere and E. Vittoz, "An analogue electronic model of Ventral Cochlear Nucleus neurons," *IEEE International Conference on Microelectronics for Neural Networks*, vol. 52-59, 1996
- [11] P. B. Manis and S. O. Marx, "Outward currents in isolated ventral cochlear nucleus neurons," *J Neurosci*, vol. 11, pp. 2865-80, 1991
- [12] G. Svirskis, et al., "Enhancement of signal-to-noise ratio and phase locking for small inputs by a low-threshold outward current in auditory neurons," *J Neurosci*, vol. 22, pp. 11019-25, 2002
- [13] E. Culurciello, R. Etienne-Cummings and K. Boahen, "A Biomorphic Digital Image Sensor," *IEEE Journal of Solid-State Circuits*, vol. 38, no. 2, pp. 281-294, 2003
- [14] K. A. Boahen, "A Burst-Mode Word-Serial Address-Event Channel," *IEEE Transactions on Circuits and Systems I*, vol. 51, no. 7, pp. 1269-1300, 2004
- [15] J. S. Rothman and P. B. Manis, "The roles potassium currents play in regulating the electrical activity of ventral cochlear nucleus neurons," *J Neurophysiol*, vol. 89, pp. 3097-113, 2003
- [16] B. Ramazan and D. Oertel, "Hyperpolarization-activated, mixed-cation current (I<sub>h</sub>) in octopus cells of the mammalian cochlear nucleus," *J Neurophysiol*, vol. 84, pp. 806-817, 2000
- [17] C. J. Sumner, et al., "A revised model of the inner-hair cell and auditory-nerve complex," *J Acoust Soc Am*, vol. 111, pp. 2178-88, 2002