Output encoding for cochlear signal analysis

Yingying Wang, Soumyajit Mandal*
Electrical Engineering and Computer Science Department, Case Western Reserve University, Cleveland, Ohio 44106
Email: sxm833@case.edu

***Abstract*—The biological inner ear, or cochlea, is an amazing sensor that performs auditory frequency analysis over an ultra-broadband frequency range of ∼20 Hz to 20 kHz with exquisite sensitivity and high energy efficiency. Electronic cochlear models, which mimic the exponentially-tapered structure of the biological inner ear using transmission lines or filter cascades, have been shown to be fast and extremely efficient spectrum analyzers at both audio and radio frequencies (RF). Here we present improved output encoding methods for such cochlea-like analyzers. We have developed neuron-like asynchronous event-generation circuits to efficiently encode cochlear outputs, including ring-oscillator-based injection-locked frequency dividers (ILFDs) that accurately encode input frequencies and phase-sensitive detectors that encode both amplitude and phase information and thereby improve frequency resolution without reducing temporal resolution.**

I. Introduction

The process of audition begins when sound waves travel down the auditory canal and vibrate the tympanic membrane, as shown in Fig. 1(a). These vibrations are coupled into the bones of the middle ear (the malleus, incus, and stapes), and then transmitted through the fluid-filled cochlea, causing fluctuations in fluid pressure (P) and volume velocity (U), and also motion of the basilar membrane (BM). The cochlea is a sophisticated signal processing system that converts BM motion into a time-varying pattern of neural excitation on the auditory nerve while consuming only ∼14 μW of power. Cochlear outputs are further processed by higher auditory centers in the brain to generate the perception of sound, resulting in exquisite sensitivity and over 120 dB of input-referred dynamic range. The cochlea can be modeled as a transmission line where shunt admittances Y model sections of the BM, while the series inductors Z model fluid coupling (see Fig. 1(b)). The values of Y and Z per unit length increase exponentially with position, which results in a frequency-dependent cutoff position for propagating waves [1]. This frequency-to-position transformation serves as the basis for cochlear frequency analysis. This spectrum analysis process is highly efficient in terms of both analysis time and hardware costs when compared to other techniques such as swept-sine (super-heterodyne), the Fast Fourier Transform (FFT), or parallel banks of independent filters [2]. In particular, the analysis time and hardware costs of the cochlea both scale linearly with N, the number of output channels.

Electronic models of cochlear mechanics rely upon a mechanical-electrical transformation in which P and U are mapped to voltage (V) and current (I). These circuits are usually approximated by a finite number of transmission line stages (series and shunt impedances) or low-pass filters [4]–[6] in on-chip implementations. The resultant cochlear transfer functions are low-pass in nature with broad peaks and very steep cutoff slopes, as shown in Fig. 1(c) and (d). They have

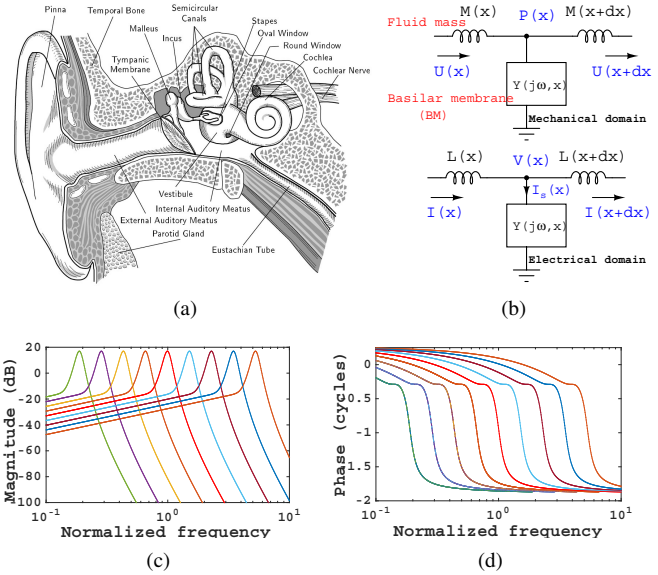

Fig. 1. (a) Anatomy of the human auditory periphery, adapted from [3]; (b) a generic spatially-varying one-dimensional transmission-line-based cochlear model in the mechanical and electrical domains; normalized cochlear transfer function (c) amplitudes and (d) phases at different positions.

been plotted on a normalized frequency scale to emphasize the fact that cochlea-like signal analysis is not limited to audio; in fact, the range of analyzed frequencies is arbitrary and can be set by appropriate scaling of element values within the circuit. We have previously used this principle to develop single-chip RF spectrum analyzers that are based on active cochlear models and operate in the 0.6-8 GHz frequency range [2]. Such ultra-broadband signal analyzers are expected to be useful for spectrum sensing in cognitive radios and other applications [7].

Previous cochlear implementations have generally used envelope detectors (EDs) that are analogous to inner hair cells to reduce the output bandwidth. The ED outputs are then sent off-chip either as analog voltages or asynchronous events (spikes) [8]–[10]. Such encoding, which is analogous to rate coding in the nervous system, only preserves low-pass-filtered versions of the output amplitudes. The resultant loss of fine time structure information present in the input severely limits the applications of both audio and RF cochlear models. One important issue is confusion between amplitude and frequency, since large inputs away from the peak frequency at a given position produce the same event rate as small inputs near the peak. Thus the output amplitude of a single stage cannot be used to decode the input. Moreover, even the outputs of multiple stages cannot resolve this ambiguity if a large signal exists at some nearby frequency. Fig. 2 shows the case when signals A with an input amplitude of 10 mV at 6.0 GHz, B

978-1-4799-7234-0/15/$31.00 ©2015 IEEE

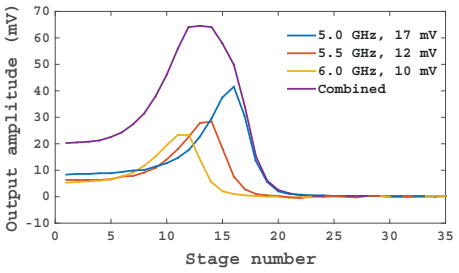

Fig. 2. Simulated cochlear output amplitudes generated by three different single-tone inputs at similar frequencies, and their combined response.

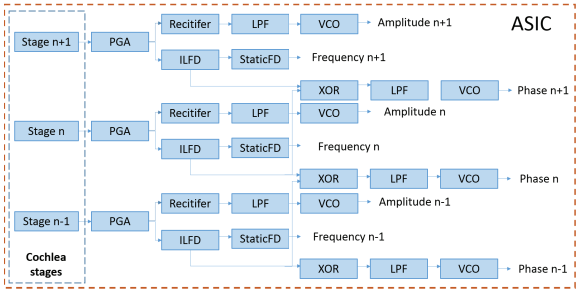

Fig. 3. Simplified block diagram of the proposed cochlea-based ultra-broadband RF spectrum analyzer chip.

(12 mV at 5.5 GHz), and C (17 mV at 5.0 GHz), are separately applied to the cochlear model. The peak output of A at stage 12 is buried under the response to B, and similarly that of B at stage 14 is buried under the response to C.

The auditory system solves this problem in several ways, such as by encoding information about input phase and frequency within phase-locked auditory nerve fibers and by cross-correlating the outputs of multiple phase-locked fibers. Phase-locking causes inter-spike intervals to cluster around integer multiples of the input period, which yields an independent estimate of the frequency. The amplitude can then be estimated unambiguously from the firing rate [11]. Furthermore, Fig. 2 shows that the first derivatives of the output amplitudes are close to zero near their peaks. Thus they provide little frequency information in regions with high sensitivity and signal-to-noise ratio (SNR). However, the magnitude of the derivative of the phase curves is maximal in such regions. Hence phase information can be used to improve frequency estimation without affecting temporal resolution [12]. In this paper we propose a novel bio-inspired cochlear encoding technique that simultaneously represents amplitudes using voltage-controlled oscillators (VCOs), frequencies using ILFDs, and local phase gradients using phase detectors (PDs) and VCOs. We will discuss integrated implementations of such encoders for power-efficient ultra-broadband RF spectrum analysis in the GHz range, but the same principles can also be used for cochlea-like signal analysis at other frequencies.

II. Cochlear Output Encoding

Fig. 3 shows the overall block diagram of the proposed cochlea-based spectrum analyzer, which has been designed in the UMC 65 nm RF-CMOS process. It includes a bidirectional cochlea model with 60 exponentially-spaced output stages, and analyzes the radio spectrum from ∼0.8 GHz to 8.3 GHz [2]. The chip also contains automatic gain control (AGC) circuits to improve dynamic range (DR), digital programmability to reduce sensitivity to component mismatches, and parallel outputs that allow events from all stages to be transmitted off-chip. The output of each cochlea stage is processed by three encoder circuits that are sensitive to amplitude, frequency, and phase shift between adjacent stages, respectively. In the amplitude branch, the cochlea output is first amplified by a three-stage programmable gain amplifier (PGA). Its envelope is then rectified, low-pass filtered, and converted into frequency by a VCO. In the frequency branch, the amplified output of the PGA injection-locks to a ring-oscillator-based frequency divider (ILFD), followed by several stages of static frequency division. The latter are implemented using current-mode logic (CML) latches. The ILFD outputs from adjacent stages are also sent into a XOR gate that acts as a PD. Another low-pass filter (LPF) and VCO then converts the XOR output into frequency, so the outputs of all three branches are encoded as frequency information. These outputs are finally multiplexed into a single wire for later decoding and analysis by an FPGA.

Amplitude encoding: Each cochlear stage output is first amplified by a wide bandwidth PGA whose gain is dynamically controlled by an AGC loop completed through the FPGA. An envelope detector [13] generates a DC voltage proportional to the amplitude of the PGA output. This voltage is then controling a differential 3-stage VCO [14] that is analogous to a rate-coding neuron, i.e., generates an asynchronous binary signal whose frequency is proportional to the analog input. By using the RF detector and the VCO, the analog amplitude information can be accurately digitized by counting the edges of the VCO output during a time period [15]. This oscillator-based quantization method requires less complex analog circuitry than voltage-mode ADCs, and is thus suitable for implementation in nanoscale CMOS processes.

Frequency encoding: Frequency dividers (FDs) are a basic but critical building block in various high-speed wireline and wireless communication systems. Conventional flip-flop based static FDs are robust and broadband but consume a lot of power in high-speed applications, while ILFDs are a low-power alternative for applications up to several GHz [16]–[18]. The ring-oscillator based ILFDs, which have wide locking ranges, are used as frequency-encoding analogs of phase-locked auditory neurons. The complementary-injection scheme shown in Fig. 4 not only increases the effect of the signal injected into the ring-oscillator but also drives both the rising and falling propagation delays, unlike conventional schemes that vary the falling propagation delays only [18]. This differential ILFD's free running frequency is mainly determined by the dimensions of the devices and the power consumption, i.e., proportional to the bias current used in the top and bottom current mirrors.

ILFD locking range depends on the power of the injected signal. The simulated locking sensitivity curve of a divide-by-3 ILFD with a free-running frequency of 2.45 GHz is shown in Fig. 5. For an input power of -21 dBm, the measured locking range is approximately 6.0 GHz to 8.3 GHz assuming default PVT (typical device corner, 1.2 V power supply, 298 K). We expect this to be large enough to compensate for fabrication

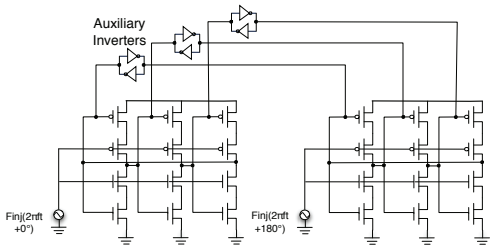

Fig. 4. Simplified schematic of the divide-by-3 differential complementary-injection-locked frequency divider (ILFD) circuit.

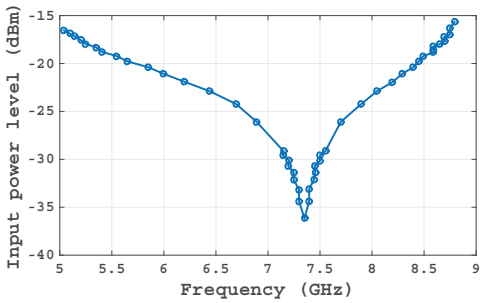

Fig. 5. Simulated input sensitivity curve for a divide-by-3 ILFD (free-running frequency at 2.45 GHz). The circuit locks when the input power level at any particular frequency exceeds the plotted value at that point.

tolerances and component mismatches.

The phases of the oscillator and the injected signals track each other in the locked state. Oscillator phase noise at offset frequencies smaller than the locking range is reduced by locking to a low-noise signal. From a time-domain perspective, the injected signal corrects the oscillator zero crossings within each period, which reduces jitter accumulation [19]. At the edges of the lock range, the injected signal cannot improve phase noise since it injects energy at a 90° phase offset, where the signal has maximum amplitude. Thus phase noise reduction decreases as the input frequency deviates from the free-running frequency, as shown in Fig. 6. The natural frequency of oscillators will incur significant error due to process variations and mismatch, which makes it difficult to rely on the phase noise reduction properties if the locking range is narrow. Our ILFD design avoids this problem because it has a wide locking range. It was simulated to have 13 dBc/Hz less phase noise in the locked state for a free-running frequency of 2.69 GHz.

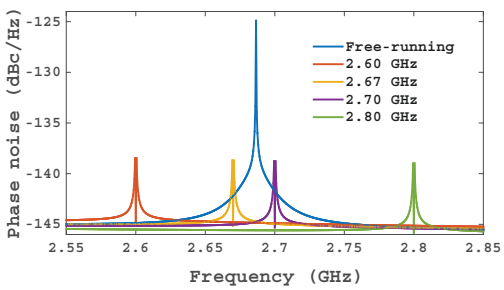

Fig. 6. Simulated phase noise for a divide-by-3 ILFD (free-running at 2.69 GHz) in the free-running and locked states for various input frequencies.

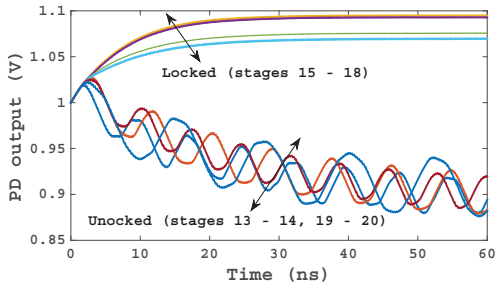

Fig. 7. Comparison between phase detector (PD) output voltages in the locked (stages 15-18) and unlocked states (stages 13-14 and 19-20).

Phase encoding: Another advantage of using an ILFD as the frequency encoder is that it preserves input phase information when locked. In other words, the phase of the locked oscillator follows that of the injection signal within its locking range, which is also the same range of phase of interest as frequency information. The phase shift across each cochlear stage, i.e., the local output phase gradient, can thus be estimated by the XOR gate shown in Fig. 3. The phase encoding circuit can also be used to estimate the locking status of the local ILFD. If the ILFD is locked, the phase shift between adjacent stages is fixed, resulting in a constant output from the XOR-LPF circuit. If the ILFD is unlocked, there is a random phase difference between adjacent stages due to the existence of inevitable component mismatches. This difference will accumulate at the LPF output and be observed as an oscillating voltage with an average value outside the normal range observed during lock, as shown in Fig. 7.

III. Simulation Results

Consider a 20 mV signal applied to a bidirectional RF cochlea containing $N = 60$ stages. High input frequencies will resonate at the initial stages (base), and low frequencies at the later stages (apex). We have developed two methods to decode cochlear outputs from measured VCO frequencies; both of these are suitable for real-time implementation on an FPGA. The first method is to generate a look-up table from the empirically-estimated ED-VCO input-output curve, and the second is to generate the theoretical polynomial equations for ED and VCO separately, and then calculate the cochlear outputs. The two methods can also be combined by adding empirical parameters into the decoding equations. Spatial responses to linearly-spaced input frequencies varying between 0.5 GHz and 8.0 GHz were decoded using the first method and are shown in Fig. 8. The decoded curves fit the actual cochlea output curves well except at the peak point, which is limited by amplifier bandwidth at high frequencies and the resolution of the look-up table.

The frequency response for an input frequency at 7 GHz is decoded and analyzed in Fig. 9, and peaks around stage 7 with an amplitude of 57 mV. The ILFD remains locked from stage 4 to stage 9, and the minimum cochlear output to ensure locking is 13 mV for default gain settings. The static frequency dividers following the ILFD remain locked over the whole range.

The fact that the theoretical phase delay across each cochlea stage has an absolute value $< 100°$ is used within a

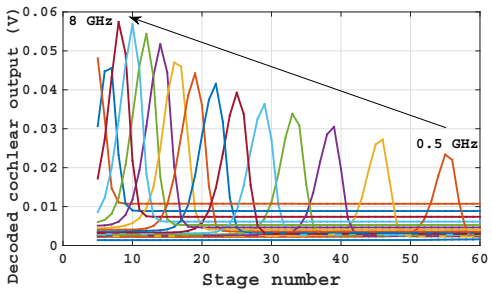

Fig. 8. Decoded cochlear output amplitudes for single-tone inputs. The input frequency was varied linearly from 0.5 GHz to 8 GHz in steps of 0.5 GHz.

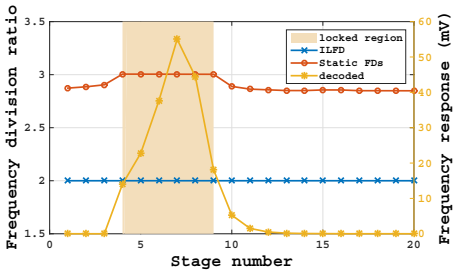

Fig. 9. Simulated (a) frequency response and (b) frequency encoding of a single-tone input at 7 GHz using a divide-by-3 ILFD followed by two static divide-by-2 stages.

verification algorithm during phase decoding. After the VCO output frequencies are decoded using empirical polynomial functions, the output goes through this algorithm, which detects and ignores non-valid data. The valid data is then converted to output phase shift values. The "simulated" and decoded phase shifts between adjacent stages are shown in Fig. 10. The two are close to each other (typical error of $\sim 5°$) in the region around stage 13 where the output amplitude is large enough for the ILFDs to be locked.

IV. Conclusion

We have described a biologically-inspired oscillator-based scheme for encoding amplitude, phase gradient, and frequency information generated by cochlea-like signal analyzers. We have also shown how the encoded outputs can be decoded and analyzed to yield signal amplitudes and frequencies. For example, on-chip VCOs generate asynchronous binary signals whose frequencies are proportional to the analog cochlear outputs. These frequencies can then be quantized off-chip by counting edges in the VCO outputs during a sampling period, thus producing multi-bit digital outputs and realizing a set of parallel time-domain ADCs. Since the VCOs produce a continuous phase output, the quantization noise of the previous sample affects that of the current sample, which results in inherent first-order quantization noise shaping. In addition, digital calibration methods can be implemented to compensate for VCO nonlinearity and increase the effective resolution of the ADCs [15]. Such methods will be explored in future work.

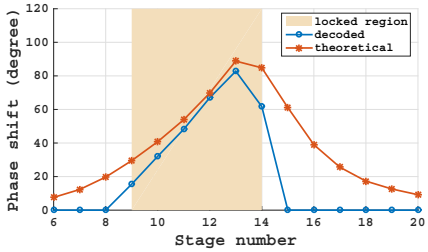

Fig. 10. "Simulated" and decoded phase shifts between adjacent cochlear outputs at an input frequency of 6 GHz.

References

[1] G. Zweig, "Finding the impedance of the organ of Corti," *J. of the Acoustical Society of America*, vol. 89, no. 3, pp. 1229–1254, 1991.

[2] S. Mandal, S. M. Zhak, and R. Sarpeshkar, "A bio-inspired active radio-frequency silicon cochlea," *IEEE J. of Solid-State Circuits*, vol. 44, no. 6, pp. 1814–1828, 2009.

[3] L. Watts, "Cochlear mechanics: Analysis and analog VLSI," Ph.D. Thesis, California Institute of Technology, Pasadena, CA, 1992.

[4] R. F. Lyon and C. Mead, "An analog electronic cochlea," *IEEE Trans. Acoustics, Speech and Sig. Proc.*, vol. 36, no. 7, pp. 1119–1134, 1988.

[5] R. Sarpeshkar, R. F. Lyon, and C. Mead, "A low-power wide-dynamic-range analog VLSI cochlea," in *Neuromorphic systems engineering*. Springer, 1998, pp. 49–103.

[6] T. J. Hamilton, C. T. Jin, A. van Schaik, and J. Tapson, "An active 2-D silicon cochlea." *IEEE Trans. Biomed. Circuits and Systems*, vol. 2, no. 1, pp. 30–43, 2008.

[7] W. J. Chappell, E. J. Naglich, C. Maxey, and A. C. Guyette, "Putting the radio in software-defined radio: Hardware developments for adaptable RF systems," *Proc. of the IEEE*, vol. 102, no. 3, pp. 307–320, 2014.

[8] S.-C. Liu, A. van Schaik, B. Minch, and T. Delbruck, "Asynchronous binaural spatial audition sensor with 2x64x4 channel output," *IEEE Trans. Biomed. Circuits and Systems*, vol. 8, no. 4, pp. 453–464, 2013.

[9] M. Newton and L. Smith, "A neurally-inspired musical instrument classification system based upon the sound onset," *Journal of the Acoustical Society of America*, vol. 131, no. 6, pp. 4785–4798, 2012.

[10] A. G. Katsiamis, E. Drakakis, and R. F. Lyon, "A biomimetic, 4.5 μw, 120+ db, log-domain cochlea channel with AGC," *IEEE J. of Solid-State Circuits*, vol. 44, no. 3, pp. 1006–1022, 2009.

[11] W. Bialek and F. Rieke, "Reliability and information transmission in spiking neurons," *Trends in Neurosciences*, vol. 15, no. 11, pp. 428–434, 1992.

[12] S. Mandal and R. Sarpeshkar, "A bio-inspired cochlear heterodyning architecture for an RF fovea," *IEEE Trans. Circuits and Systems I*, vol. 58, no. 7, pp. 1647–1660, July 2011.

[13] A. Valdes-Garcia, R. Venkatasubramanian, R. Srinivasan, J. Silva-Martinez, and E. Sanchez-Sinencio, "A CMOS RF RMS detector for built-in testing of wireless transceivers," in *IEEE VLSI Test Symposium*, 2005, pp. 249–254.

[14] G. Jovanovic, M. Stojcev, and Z. Stamenkovic, "A CMOS voltage controlled ring oscillator with improved frequency stability," *Applied Mathematics, Informatics and Mechanics (AMIM)*, vol. 2, no. 1, 2011.

[15] J. Kim, T.-K. Jang, Y.-G. Yoon, and S. Cho, "Analysis and design of voltage-controlled oscillator based analog-to-digital converter," *IEEE Trans. Circuits and Systems I*, vol. 57, no. 1, pp. 18–30, 2010.

[16] S. Verma, H. R. Rategh, and T. H. Lee, "A unified model for injection-locked frequency dividers," *IEEE J. of Solid-State Circuits*, vol. 38, no. 6, pp. 1015–1027, 2003.

[17] J.-C. Chien and L.-H. Lu, "Analysis and design of wideband injection-locked ring oscillators with multiple-input injection," *IEEE J. of Solid-State Circuits*, vol. 42, no. 9, pp. 1906–1915, 2007.

[18] Y.-C. Lo, H.-P. Chen, J. Silva-Martinez, and S. Hoyos, "A 1.8 V, sub-mW, over 100% locking range, divide-by-3 and 7 complementary-injection-locked 4 GHz frequency divider," in *IEEE Custom Integrated Circuits Conference (CICC)*. IEEE, 2009, pp. 259–262.

[19] B. Razavi, "A study of injection locking and pulling in oscillators," *IEEE J. of Solid-State Circuits*, vol. 39, no. 9, pp. 1415–1424, 2004.