

A 1.0-8.3 GHz Cochlea-Based Real-Time Spectrum Analyzer With Δ - Σ -Modulated Digital Outputs

Yingying Wang^{1b}, *Member, IEEE*, Gihan J. Mendis^{1b}, Jin Wei-Kocsis^{1b}, *Member, IEEE*,
Arjuna Madanayake^{1b}, *Member, IEEE*, and Soumyajit Mandal^{1b}, *Senior Member, IEEE*

Abstract—The biological inner ear, or cochlea, is a sophisticated signal processing system that performs spectrum 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 bidirectional transmission lines or filter cascades, act as fast and hardware-efficient spectrum analyzers at both audio and radio frequencies. This paper describes a cochlea-based digitally-programmable single-chip radio frequency (RF) spectrum analyzer in 65 nm CMOS. This “RF cochlea” chip includes a transmission-line active cochlear model with 50 parallel exponentially-spaced stages that analyzes the radio spectrum from 1.0-8.3 GHz. The outputs of all stages are encoded in parallel as delta-sigma (Δ - Σ) modulated digital signals for real-time demodulation and analysis by a digital back-end processor. The chip consumes 418 mW and typically generates ~ 1 GS/s of total data at an ENOB of 5-6 bits. An artificial intelligence (AI)-driven single-channel cognitive radio (CR) receiver based on the RF cochlea has also been implemented and tested.

Index Terms—Cognitive radio, radio frequency (RF) cochlea, bio-inspired, spectrum analysis, modulation recognition.

I. INTRODUCTION

RADIO frequency (RF) spectrum is a valuable but tightly regulated resource due to its unique and important role in wireless communications. The commercially interesting RF and microwave spectrum, which covers the frequency range from 100 MHz to 10 GHz, is becoming increasingly crowded. Moreover, wireless users want ever-higher quality of service, particularly increased reliability and higher data rates. Thus, there is an urgent need for higher spectral efficiency. Cognitive radio (CR) is a promising candidate for improving spectrum

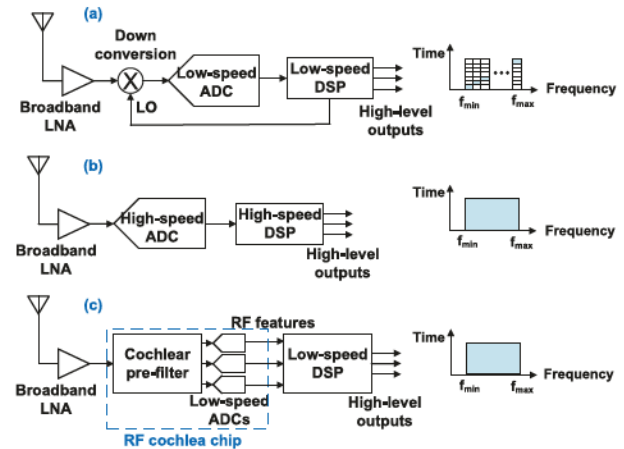


Fig. 1. (a) Conventional narrow-band, (b) conventional broad-band, and (c) proposed CR receiver front-end architectures. The plots on the right show typical time-frequency coverage patterns for each architecture.

utilization in next generation wireless networks [1], [2]. CR adds intelligence to traditional wireless communication systems by dynamically adapting the operating parameters, such as transmit power, carrier frequency and modulation strategy, based on the local RF environment. In particular, CR systems actively search for spectrum holes or white spaces, i.e., time-frequency locations that are free of interfering signals and are thus suitable for reliable communications [3]–[5].

Given the need to find white spaces, wideband real-time spectrum sensing is a crucial requirement for future CR networks. Most implementations use one of various multi-resolution spectrum sensing (MRSS) algorithms [6], [7]. Two conventional single-antenna CR receiver front-end architectures incorporating spectrum sensing are shown in Figs. 1(a)–(b). Fig. 1(a) shows a conventional heterodyne-based CR with down-converting channels, in which the local oscillator (LO) is controlled by a software-defined radio (SDR). This architecture has limited real-time bandwidth and so can only monitor one narrow band at a time, which results in low overall spectrum coverage. By contrast, Fig. 1(b) is based on directly digitizing the amplified RF waveform from the antenna, followed by high-speed digital signal processing (DSP) to extract information such as spectral occupancy and modulation type. This architecture provides full coverage of a broad band, but is not energy-efficient since a large fraction of the digitized data is subsequently discarded.

Here we describe a single-chip real-time RF spectrum analyzer for CR applications that is based on a model of

Manuscript received December 18, 2019; revised February 20, 2020 and March 30, 2020; accepted April 21, 2020. Date of publication May 6, 2020; date of current version September 2, 2020. This project was sponsored by the National Science Foundation (NSF) under Grant CCF-1525162. This article was recommended by Associate Editor B. Zhao. (*Corresponding author: Yingying Wang.*)

Yingying Wang and Soumyajit Mandal are with the Department of Electrical, Computer, and Systems Engineering, Case Western Reserve University, Cleveland, OH 44106 USA (e-mail: yxw788@case.edu; sxm833@case.edu).

Gihan J. Mendis and Jin Wei-Kocsis are with the Department of Computer and Information Technology, Purdue University, West Lafayette, IN 47907 USA (e-mail: kocsis0@purdue.edu).

Arjuna Madanayake is with the Department of Electrical and Computer Engineering, Florida International University, Miami, FL 33174 USA (e-mail: amadanay@fiu.edu).

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TCSI.2020.2990364

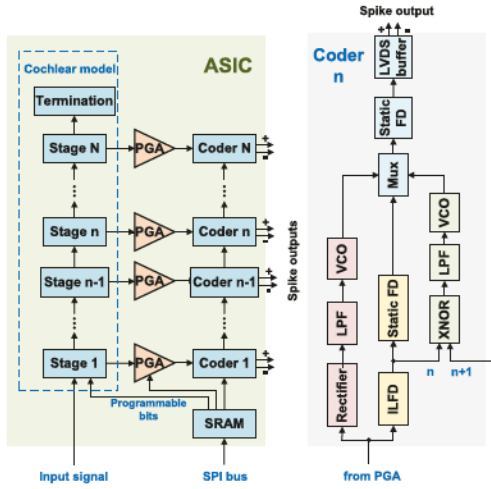


Fig. 2. Simplified block diagram of the proposed cochlea-based ultra-broadband RF spectrum analyzer chip.

the mammalian cochlea (inner ear). This ultra-broadband “RF cochlea” chip improves system-level energy efficiency by performing a significant amount of low-power analog processing (pre-filtering and signal detection) prior to distributed low-data-rate analog-to-digital conversion. In particular, the on-chip cochlear pre-filter extracts a set of lower-bandwidth RF features from wideband inputs (including amplitude, phase, and frequency information for spectral occupancy and modulation recognition), thus reducing DSP speed and power requirements as shown in Fig. 1(c). In other words, the proposed architecture saves power at the system level by using analog pre-processing to extract meaningful features prior to digitization [8].

This paper describes both a complete working RF cochlea chip and its integration within an AI-driven CR receiver of the type shown in Fig. 1(c). It is organized as follows. In Sections II and III, we describe the system- and circuit-level design of the chip, respectively. Simulation and measurement results on both the chip and a CR receiver based on it are presented in Section IV. Finally, Section V concludes the paper.

II. SYSTEM DESIGN

A. Overall Chip Architecture

Fig. 2 shows a block diagram of the chip. It includes a bidirectional cochlear model that will be described in the next section. This model consists of $N = 50$ exponentially-scaled stages that analyze the radio spectrum from 1.0-8.3 GHz into N frequency bins with approximately constant fractional bandwidth. Thus, these outputs resemble those generated by a constant- Q filter bank. Each output is processed by three encoder circuits that are sensitive to signal amplitude, frequency, and phase shift between adjacent stages, respectively.

Each amplitude encoder first amplifies the output of a cochlear stage using a two-stage programmable gain amplifier (PGA). It then rectifies, low-pass filters, and converts the signal into frequency using a voltage-controlled oscillator (VCO). The frequency encoder first locks a ring-oscillator-based

injection-locked frequency divider (ILFD) to the amplified signal, followed by several stages of static frequency division implemented using current-mode logic latches. The phase encoder feeds the ILFD outputs from adjacent stages into a XNOR gate that acts as a phase detector (PD); another low-pass filter and VCO then converts the PD output into frequency. The outputs of all three encoders are thus represented as single-bit asynchronous digital sequences. These sequences are finally multiplexed into a single low-voltage differential signaling (LVDS) signal for off-chip decoding and analysis.

Note that the encoded digital outputs are available in parallel from all 50 stages, unlike earlier designs where analog voltages were serially scanned off-chip [9]. This change greatly increases the rate at which information can be transmitted off-chip for real-time digital signal processing. In addition, the new design also adds a large amount of digital programmability via a standard serial peripheral interface (SPI).

B. Cochlear Signal Analysis

The process of audition begins when sound waves travel down the auditory canal and vibrate the tympanic membrane, as shown in Fig. 3(a). These vibrations are transduced into motion of the basilar membrane (BM) within the inner ear (cochlea). The cochlea contains a sophisticated signal processing system that converts BM motion into a time-varying pattern of neural excitation on the auditory nerve while consuming only $\sim 14 \mu\text{W}$ of power [8]. 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. To a good approximation, the cochlea can be electrically modeled as a transmission line where shunt admittances Y model sections of the BM, while the series inductors Z model fluid coupling (see Fig. 3(b)). The values of Y and Z per unit length increase exponentially with position, which results in a frequency-dependent cutoff position for waves propagating on the BM [10]. This frequency-to-position transformation is the basis for cochlear frequency analysis.

Cochlear models are usually implemented by spatial discretization [13], [14], i.e., approximating the transmission line shown in Fig. 3(b) with a finite number of stages. These stages may consist of i) bidirectional series and shunt impedances, or ii) unidirectional low-pass filters. In either case, the resultant cochlear transfer functions are low-pass with broad peaks and very steep cutoff slopes (Fig. 3(c)-(d)). These functions have 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. Thus, electronic cochlear models that mimic the exponentially-tapered structure of the inner ear act as fast and hardware-efficient spectrum analyzers at both audio frequencies [13], [15]–[18] and RF [9], [19]–[22]. In particular, we have previously used this principle to develop single-chip RF spectrum analyzers (“RF cochleas”) that are based on active bidirectional or unidirectional cochlear models and operate in the 0.6-8 GHz frequency range [9], [21].

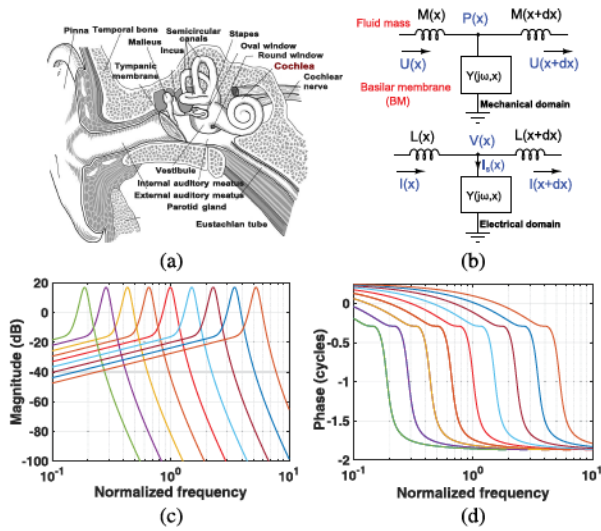


Fig. 3. (a) Anatomy of the human auditory periphery, adapted from [11]; (b) a generic spatially-varying one-dimensional transmission-line-based cochlear model in the mechanical and electrical domains. Normalized (c) amplitudes and (d) phases of cochlear transfer functions to exponentially-spaced normalized input frequencies varying between 10^{-1} and 10^1 [12].

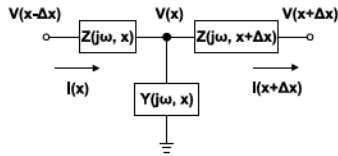


Fig. 4. A generic spatially-varying one-dimensional transmission line, with series impedance represented by Z and shunt admittance by Y .

C. Implementation of the Cochlear Model

The usual transformation between mechanical and electrical domains used to model the cochlea relies on mapping fluid pressure P to voltage V and volume velocity U to current I , respectively (i.e., $(P, U) \rightarrow (V, I)$); this results in electrical models of the form shown in Fig. 3(b). Here we use the dual transformation $(P, U) \rightarrow (I, V)$, which results in an electrical model that is more amenable to on-chip integration at microwave frequencies [9]. The equations for voltage V (corresponding to U) and current I (corresponding to P) on the resulting spatially-varying transmission line shown in Fig. 4 in sinusoidal steady-state are given by

$$\frac{dV}{dx} = -Z(\omega, x)I \quad \text{and} \quad \frac{dI}{dx} = -Y(\omega, x)V, \quad (1)$$

where $Z(\omega, x)$ and $Y(\omega, x)$ are the impedance and admittance per unit length of the line. As in the biological cochlea, we assume that the spatially-varying characteristic frequency ω_c of the structure decreases exponentially with position x , i.e., as $\omega_c(x) = \omega_c(0) \exp(-x/l)$ where l is a characteristic length. In this case, the dependence of Z and Y on both frequency (ω) and spatial position (x) can be replaced by dependence on a single normalized frequency variable $s_n = j\omega/\omega_c(x)$ [10].

Moreover, for circuit implementations, the continuously-varying line impedance and admittance (Z and Y) are i) lumped into a finite number of stages, each of length Δx ; and ii) converted into dimensionless quantities Z_n and Y_n by

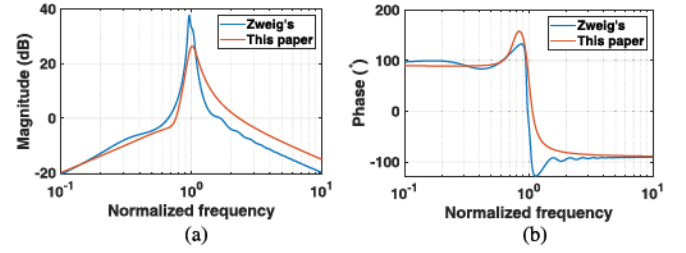


Fig. 5. Comparison of (a) magnitude and (b) phase of the series impedance Z_n between experimental data from Zweig [10] and this paper.

a fixed impedance Z_0 , as follows:

$$Z_n = (\Delta x)Z(s_n)/Z_0, \quad Y_n = (\Delta x)Y(s_n)Z_0. \quad (2)$$

Our previous work [9], [23] has shown that experimental data from alive mammalian cochleas is well-modeled by assuming that i) the shunt admittance Y_n is simply a capacitor, and ii) the series impedance Z_n is a rational function with three zeros and four poles. In particular, we use

$$Y_n(s_n) = s_n Q_{line}, \quad Z_n(s_n) = \frac{s_n(s_n^2 + \frac{\mu}{Q}s_n + \mu^2)}{\mu^2(s_n^2 + 2ds_n + 1)^2}, \quad (3)$$

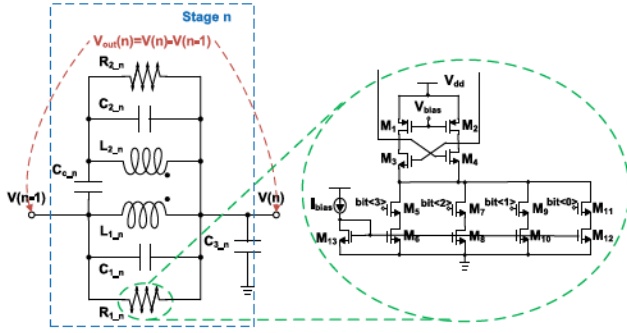
where μ , Q and d are constants, and $Q_{line} = \omega_c(0)C_0Z_0$ where C_0 is the capacitance corresponding to the admittance of the first stage. Physically, the dimensionless parameter Q_{line} is a measure of spatial discretization; it should be $\ll 1$ to ensure that wave propagation on the cochlea is not affected when we approximate it with a finite number of stages.

We optimized the parameter values in (3) to match experimental data in [10]. In particular, for the series impedance we found that $d = 0.1$, $\mu = 0.76$, and $Q = 3.8$ provide a good fit; the resultant form of Z_n is shown in Fig. 5(a) and (b). Note that this set of parameters describes a *active* cochlear model in the sense that $Z_n(s_n)$ cannot be realized using only passive components. Physically, each stage in such models contains a positive feedback loop that feeds energy into the propagating wave (modeling the action of outer hair cells in the biological cochlea); this results in higher sensitivity and frequency resolution than passive cochlea-like filter banks [9].

We also selected $\omega_c(0) = 2\pi \times 10^{10}$ rad/s to set the upper frequency limit to ~ 10 GHz, $Q_{line} = 0.08$ to minimize the effects of spatial discretization, and $Z_0 = 56.6 \Omega$. The choice of Z_0 is determined by a trade-off between three factors: i) power consumption, which decreases with Z_0 since the transconductances of transistors within the series impedance Z scale as $g_m \propto 1/Z_0$; ii) layout area, which increases with Z_0 since the inductances within Z scale as $L \propto Z_0$; and iii) input impedance, which is proportional to Z_0 and should be in the 20-200 Ω range for compatibility with RF components and test equipment. In particular, the input impedance at frequencies much smaller than $\omega_c(0)$ is real and given by

$$Z_{in} = Z_0 \sqrt{Z_n/Y_n} = Z_0 / \sqrt{Q_{line}}. \quad (4)$$

In this design, we set $Z_{in} = 200 \Omega$, which is at the upper end of our range, to minimize power consumption.

Fig. 6. Cochlea circuit diagram for stage n .

The schematic of a single cochlear stage is shown in Fig. 6. It is easily verified that the coupled resonators have an impedance of the right form for realizing Z_n , while the admittance of the shunt capacitor matches Y_n . Both these normalized variables are scaled by Z_0 or $1/Z_0$ to convert them back to real impedances or admittances, respectively. Finally, the desired exponential decrease of center frequency with position is accomplished by scaling all inductor and capacitor values exponentially with stage number n . Specifically, we let $L_1(n)$, $L_2(n)$, $C_1(n)$, $C_2(n)$, $C_c(n)$ scale as $\exp(n/N_{nat})$ where $N_{nat} = l/(\Delta x)$ sets the rate of spatial scaling, while keeping the resistors $R_1(n)$, $R_2(n)$ fixed [11], [24].

Finding component values for the various inductors, capacitors, and resistors that implement $Z_n(s_n)$ in (3) is a network synthesis problem with multiple possible solutions. We solved it by constrained minimization of a cost function defined in Mathematica, as described in our earlier work [9]. The following optimized parameters were obtained and used for on-chip implementation: $L_1(0) = 0.90$ nH, $L_2(0) = 1.02$ nH, $k = 0.780$, $C_1(0) = 61.5$ fF, $C_2(0) = 109.0$ fF, $C_c(0) = 1.38$ pF, $C_3(0) = C_0 = 22.5$ fF, $R_1 = -71.5 \Omega$, $R_2 = 53.6 \Omega$, and $N_{nat} = 24$. Note that R_1 is negative, as expected for an active cochlear model, and must be implemented using active elements as shown in Fig. 6. A cross-coupled NMOS pair $M_{3,4}$ sets the small-signal negative resistance to $R_1 = -2/g_m$, where the transconductance g_m can be digitally adjusted via a 4-bit DAC on the bias current I_{bias} . The DC voltage on the cochlear transmission line is determined by the bias voltage V_{bias} applied at the gates of the PMOS transistors $M_{1,2}$; this is automatically set by a slow global feedback loop (not shown) that uses an operational transconductance amplifier (OTA) to sense the actual DC voltage, compare it to a reference $V_{ref} \approx 0.7V_{DD}$, and set V_{ref} accordingly.

In practice we have to account for the significant parasitic capacitances of L_1 , L_2 , and the transistors used to implement R_1 . Fortunately, the proposed structure allows these unwanted components to be lumped into C_1 , C_2 , C_c , and C_3 . Backward wave propagation due to reflections from the low-frequency end of the transmission line can also be minimized by adding a passive termination impedance Z_t that consists of a parallel RC circuit, i.e., $Z_t = R_t || 1/(sC_t)$ [9]. The values of $R_t \approx Z_{in}$ and C_t are selected to provide a good impedance match up to the lowest analysis frequency, i.e., $\omega_c(N)$ where there are a total of N cochlear stages. Finally, the output of each stage is

measured across Z_n :

$$V_{out}(n) = V(n) - V(n-1) \approx (\Delta x) \times (dV/dx). \quad (5)$$

For a wave of frequency ω , the “best” position n_{best} at which the amplitude of $V_{out}(n)$ reaches its maximum has a characteristic frequency that approximately matches ω , i.e., $\omega_c(n_{best}) \approx \omega$. Since $\omega_c(n) = \omega_c(0) \exp(-n/N_{nat})$, we get

$$n_{best} \approx N_{nat} \ln(\omega/\omega_c(0)). \quad (6)$$

This logarithmic mapping between frequencies and the locations of spatial peaks is key to cochlear signal analysis [13].

The maximum signal-to-noise ratio (SNR) of the outputs is

$$SNR_{max} \approx (V_L^2/2)/\overline{v_{n,out}^2}, \quad (7)$$

where V_L is the linear range of the negative resistance circuit, and $\overline{v_{n,out}^2}$ is the total output noise. The former depends on transistor overdrive voltage (and thus on bias current I_{bias}), while the latter depends both on the parameters of the cochlear model (d , μ , Q , Q_{line} , and N_{nat}) and the stage number n . In general, $\overline{v_{n,out}^2}$ quickly increases with n as noise from the earlier stages accumulates, then gradually decreases as the output bandwidth shrinks along with $\omega_c(n)$. A more detailed analysis is available in our previous work [21]. Assuming typical parameter values of $V_L = 0.2$ V and $\overline{v_{n,out}^2} = 300 \mu V_{rms}$ (over an integration bandwidth from 1 kHz to 10 GHz), we get $SNR_{max} \approx 54$ dB. In addition, distributed gain compression occurs for larger signal amplitudes, resulting in a dynamic range (DR) that exceeds SNR_{max} and generally approaches ~ 65 dB. In this context, “distributed” means that the gain compression observed at one stage is the accumulated result of compression in the $N_{nat} \gg 1$ previous stages of the exponentially-tapered transmission line structure.

D. Signal Encoding Method

Previous cochlear implementations have generally used envelope detectors (EDs), which are analogous to inner hair cells, to reduce the output bandwidth. ED outputs are typically sent off-chip either as analog voltages or asynchronous events (“spikes”) [25]. Such encoding 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 the cochlear transfer functions are low-pass in nature, large inputs away from the peak frequency at a given position can 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.

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 [26]. Furthermore, phase information can be used to improve frequency estimation without affecting temporal resolution [21]. Phase information preserved by the phase-locked outputs of two independent cochlea chips can also be extracted by a cross-correlation network and used to localize sources, in a manner analogous to binaural hearing [27]. Hence novel circuits for encoding both frequency (using ILFDs) and local phase gradient (using PDs and VCOs) are implemented in this design, as shown in Fig. 2. These circuits generate bio-inspired asynchronous parallel digital outputs (spikes) from the cochlear model.

III. CIRCUIT DESIGN

A. Automated Design of Integrated Transformers

High-quality on-chip transformers are critical for realizing the cochlear stages shown in Fig. 6, since finite quality factor (Q) causes propagating waves to attenuate before reaching their characteristic or “best” positions, thus degrading input-referred noise. On-chip transformers based on regular metal polygons [28], [29] can realize a broad range of winding inductance values (typically 0.5-50 nH), but are limited to peak Q values of 5-20. A few such structures can be hand-optimized, but this becomes unfeasible for the cochlea chip, which requires $N = 50$ of them. In this section we describe a closed-loop design algorithm for integrated inductors and transformers that automates this process by combining the intuitive and fast results of theoretical calculations with accurate ones from finite-element electromagnetic (EM) simulations. It was found to converge within a few iterations in all cases.

Fig. 7 shows the overall block diagram of the proposed algorithm, in which a gradient-based optimization function sets the inductor dimensions that provide the desired winding inductance. First, we define a cost function that incorporates the synthesis goal (defined inductance value) subject to various design constraints (e.g. layout area, minimum Q over a specified frequency range, etc). The algorithm then generates a group of layout parameters (e.g., outer dimension d_{out} , metal spacing s , and metal width w) that satisfies the constraints. A MATLAB script transforms these parameters into the position coordinates of each metal segment of the inductor, allowing the geometry to be automatically imported into the EM simulator, which in our case was Sonnet (Sonnet Software, Syracuse, NY). The latter is then used to obtain a more accurate value for the inductance, which in turn allows us to calculate the error between the designated and simulated inductance values. This error is minimized by iterating the optimization process until it converges (see Fig. 7). The optimized geometry is then transferred into the circuit simulator (Cadence Virtuoso).

The Q of on-chip inductors is limited by a combination of series resistance and substrate losses. The former increases with frequency because of the skin and proximity effects. We model the resulting frequency-dependent resistance as follows:

$$R(\omega) \approx \begin{cases} R_{dc}[1 + \frac{1}{10}(\frac{\omega}{\omega_{crit}})^2], & \text{if } \delta \geq r \\ R_{dc}[\frac{1}{1-(\delta/r)^2} + \frac{1}{10}(\frac{\omega}{\omega_{crit}})^2], & \text{if } \delta < r \end{cases} \quad (8)$$

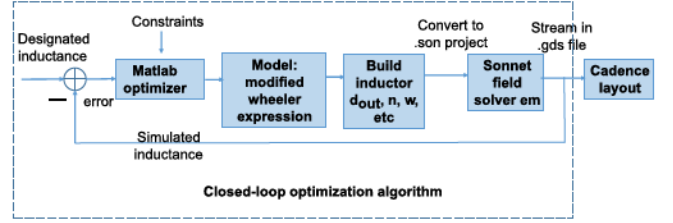


Fig. 7. Block diagram of the proposed closed-loop optimization algorithm.

where $\delta = \sqrt{\frac{2}{\omega\mu\sigma}}$ is the skin depth, σ is the conductivity, t is the thickness, μ is the permeability of free space ($\mu = 4\pi \times 10^{-7}$ H/m). The estimated series resistance shown in (8) is used as one of the constraints for the optimization algorithm.

The proposed transformer geometry uses vertically stacked octagonal windings to maximize inductance density and coupling constant k . Since the value of k affects the cochlear series impedance Z_n , it must also be set to a desired value ($k_{target} = 0.780$). A specific k can be realized by laterally displacing the centers of the primary and secondary inductors L_1 and L_2 ; the amount of displacement d_{xy} along each axis is determined by a similar closed-loop optimization algorithm after the inductor parameters have been fixed. The resulting value of k decreases linearly as d_{xy} increases before becoming negligible when the displacement becomes larger than the outer dimension d_{out} of the smaller inductor. Thus, we can analytically model k between two similarly-sized inductors as

$$k(d_{xy}) \approx \begin{cases} k(0)(1 - \frac{d_{xy}}{d_{out}}), & \text{if } d_{xy} < d_{out} \\ 0, & \text{if } d_{xy} \geq d_{out} \end{cases} \quad (9)$$

Here $k(0)$ was found to increase gradually with inductor size, from ~ 0.75 (stage 1) to ~ 0.90 (stage 50). Eqn. (9) provides a good starting point for the optimization algorithm. It also gives us some insight when setting the target k for the $N = 50$ stages. Due to the larger inductances for the later stages (i.e. > 2 nH), we increased the k_{target} for stages 11 $\sim$ 50 from 0.780 to 0.842 to save layout area, and re-optimized the other component values in Z_n accordingly.

This design uses the UMC 65 nm CMOS process, which has 8 Cu layers and 1 thick Al layer for pad connections. The algorithm generated 50 exponentially-scaled octagonal transformers to implement the cochlear model, with top and bottom windings on the Al and ultra-thick top Cu layers, respectively. The design goal was to maximize Q at a typical frequency $\sqrt{\omega_c(0)\omega_c(n)}$ subject to constraints on i) inductance values (within $\pm 1\%$ of the target at the best frequency $\omega_c(n)$), ii) total layout area (kept fixed to simplify layout).

As an example, the layout view of the transformer for stage 36 is shown in Fig. 8(a). Fig. 8(b) lists the simulated primary and secondary inductances (L_1 and L_2 , respectively) for all 50 transformers after optimization, and compares them with the desired (i.e., theoretical) values. The simulated and desired values are in good agreement, with typical errors of $\sim 0.5\%$. Similarly, Figs. 8(c) and (d) shows the typical Q values and coupling factors of the optimized transformers, respectively. A broadband S-parameter model of each transformer was extracted from Sonnet and used for chip-level simulations.

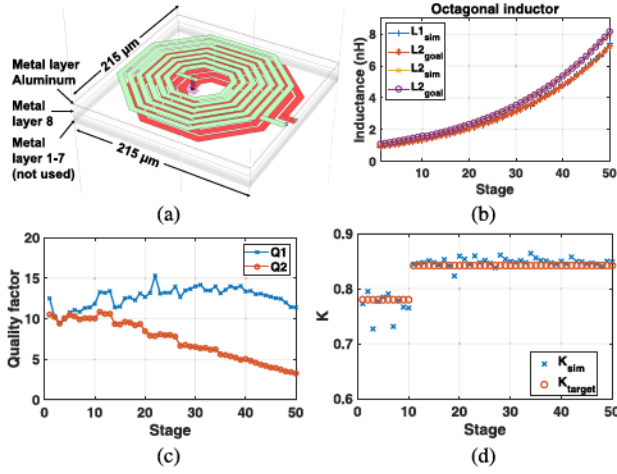


Fig. 8. (a) 3-D layout view for the octagonal transformer designed for stage 36. (b)-(d) Summary of transformer properties for stages 1-50: (b) desired and simulated inductance, (c) simulated Q at $\omega = \sqrt{\omega_c(0)\omega_c(n)}$, and (d) desired and simulated coupling coefficient.

Unwanted coupling between transformers in adjacent stages was also modeled using EM simulations. If we only consider coupling from horizontally-adjacent stages, the effects on cochlear transfer functions are relatively minor. However, if we also consider coupling from vertically-adjacent stages, some high-frequency feed-through occurs (at the -40 dB level) in the outputs of later stages. Future work will focus on improving the layout to decrease such mutual coupling effects.

B. Amplitude Encoding

Each cochlear stage output $V_{out}(n)$ is first amplified by a wideband PGA, as shown in Fig. 9(a)). The PGA uses a cascoded differential pair with resistive loads and shunt peaking to improve bandwidth. The shunt peaking inductor is a standard design from the foundry-supplied process libraries; it uses metal layers 1 $\sim$ 6 to minimize layout area and has an inductance of 18.3 nH. The PGA voltage gain A_{PGA} can be digitally set using a bias current DAC. An ED (based on MOS diodes) generates a DC voltage proportional to the amplitude of the PGA output. The primary role of the PGA is to improve the output DR by reducing the input-referred dead zone V_D of the ED. In particular, V_D decreases from $\phi_T/\kappa \approx 35$ mV, where $\phi_T = kT/q$ is the thermal voltage, to $\phi_T/(\kappa A_{PGA})$.

The ED output controls the frequency of a differential 3-stage ring oscillator [30] (shown in Fig. 9(b)) that is analogous to a rate-coding neuron, i.e., generates an asynchronous binary signal whose frequency is proportional to the analog input. Equivalently, the system implements a VCO-based quantizer (as shown in Fig. 9(c)), which behaves as a first-order Δ - Σ modulator [31]. Thus the ED and VCO accurately encode oversampled analog amplitude information, which can be digitized off-chip by counting VCO edges over periods of a user-defined sampling clock (of frequency f_{clk}). These values are then latched and decimation filtered. This oscillator-based quantization method requires less complex analog circuitry than voltage-mode ADCs, and also scales better to nanometer processes. The signal-to-noise-and-distortion ratio (SNDR)

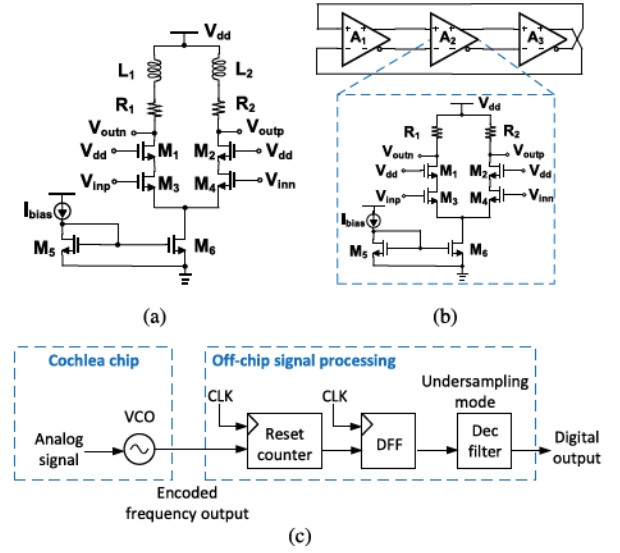


Fig. 9. Circuits used for amplitude encoding: (a) PGA, and (b) VCO. (c) Block diagram of a VCO-based quantizer.

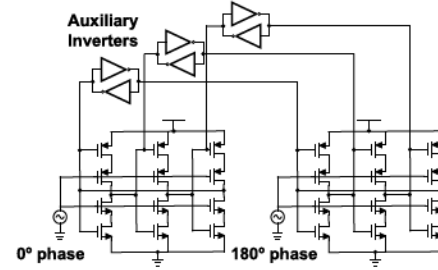


Fig. 10. Simplified schematic of the divide-by-3 differential complementary-injection-locked frequency divider (ILFD) circuit.

and effective number of bits (ENOB) of the quantizer are mainly limited by VCO jitter and nonlinearity. Simulations show that the proposed VCO has a free-running frequency range of 0.96 $\sim$ 1.76 GHz. Its simulated phase noise over bandwidth while free-running at 1.21 GHz is 0.140 rad, resulting in a total rms jitter of ~ 18.4 ps and an expected ENOB of ~ 6 bits for a typical $f_{clk} = 20$ MHz.

C. Frequency Encoding

ILFDs provide a low-power alternative to conventional flip-flop based static frequency dividers (FDs) for applications up to several GHz [32]–[34]. We use ring-oscillator based divide-by-3 ILFDs, which have wide locking ranges, as frequency-encoding analogs of phase-locked auditory neurons. Our differential ILFD design uses a complementary-injection scheme, shown in Fig. 10, that i) increases the effect of the injected signal on the ring oscillator; and ii) drives both the rising and falling propagation delays, unlike conventional schemes that vary falling delays only [34]. The ILFD's free-running frequency is mainly determined by device dimensions 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 amplitude of the injected signal. The simulated locking sensitivity curve of a divide-by-3 ILFD with a free-running frequency of 2.45 GHz

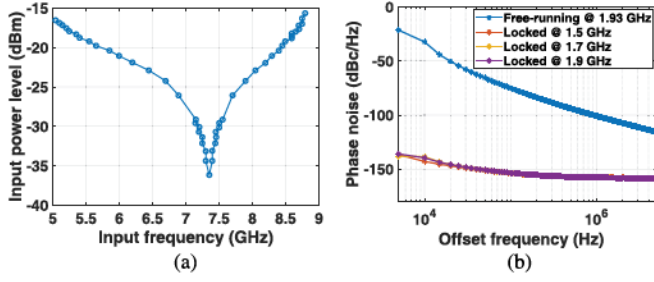


Fig. 11. (a) Simulated input sensitivity curve for a divide-by-3 ILFD (free-running frequency of 2.45 GHz). The circuit locks when the input power level at any particular frequency exceeds the plotted value at that point. (b) Simulated phase noise for a divide-by-3 ILFD (free-running at 1.93 GHz) in the free-running and locked states for various input frequencies.

is shown in Fig. 11(a). For an input power of -21 dBm, the locking range is approximately 32% (6.0–8.3 GHz) assuming default PVT (typical process corner, $V_{DD} = 1.2$ V, 298 K).

The phase of an ILFD in the locked state tracks that of the injected signals. As in a PLL, output phase noise at offset frequencies smaller than the locking range is determined by the input signal. From a time-domain viewpoint, the injected signal corrects the oscillator zero crossings within each period, which reduces jitter accumulation [35]. 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. 11(b).

An ILFD acts like a tuned detector with a detection bandwidth that increases with signal amplitude (see Fig. 11 for an example). Thus, for optimum sensitivity at low input amplitudes, the ILFD free-running frequencies should match the exponentially-varying characteristic frequency function $\omega_c(n) = \omega_c(0) \exp(-n/N_{nat})$ of the cochlea. In our case, such a match must be maintained across a 8:1 range in $\omega_c(n)$ (i.e., for $N = 50$ stages) in the presence of PVT variations and mismatch. This is ensured by using two-level frequency tuning: i) coarse tuning, in which the 50 ILFDs are divided into 4 “banks”, each with a different nominal free-running frequency; and ii) fine tuning, in which each ILFD’s free-running frequency is programmed via a 3-bit current DAC.

D. Phase Encoding

Another advantage of using an ILFD 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. The phase shift across each cochlear stage $\Delta\phi$ (and thus the local output phase gradient) can then be estimated by a XNOR gate, which acts as a PD. The PD output is averaged by a OTA-based low-pass filter (LPF). Note that a XNOR is used instead of its inverse (a XOR) for convenience. In particular, $\Delta\phi$ is limited to the range $[0, 80^\circ]$, so XNOR and XOR outputs tend to remain near V_{DD} and ground, respectively; the former is better-matched to the input voltage range of the LPF. Simulated phase encoding function at the output of the VCO is shown in Fig. 12(a) for

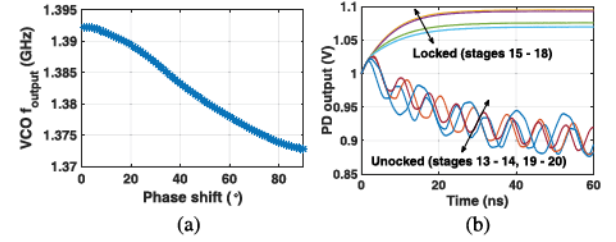


Fig. 12. (a) Simulated VCO output frequency when sweeping input signal phase shift at 2.0 GHz; (b) comparison between PD output voltages in the locked (stages 15–18) and unlocked states (stages 13–14 and 19–20).

TABLE I
SUMMARY OF CHIP DESIGN PARAMETERS

Block	SPI bits	Range	Default	Notes
Resistor R_1	4	0.5–7.5 mA	2 mA	
PGA	2	200–400 μ A	280 μ A	Max gain: 12 dB BW: 9.0 GHz
ILFD	3	150–360 μ A	240 μ A	Bank A
		60–200 μ A	120 μ A	Bank B
		30–100 μ A	60 μ A	Bank CD
		5–40 μ A	20 μ A	Bank EF
LPF	2	5–40 μ A	20 μ A	
Multiplexer	2			Amplitude, Frequency, Phase
Static FD	1	$/4 - /8$	$/8$	Bypass
Total	14			

an input frequency of 2.0 GHz. The encoding is approximately linear over the expected phase shift range of $\Delta\phi = [0, 80^\circ]$.

The phase encoding circuit can also be used to estimate the locking status of the local ILFDs. If the ILFDs of adjacent stages are locked, the phase shift between them is fixed, resulting in a constant output from the XNOR-LPF circuit. If the ILFDs are unlocked, their free-running frequencies are mismatched, which results in a random time-varying phase difference between adjacent stages. This results in an oscillating output voltage with an average value outside the normal range observed during lock, as shown in Fig. 12(b).

E. Output Multiplexer and Buffer

The outputs of the frequency-encoding paths described above are fed into a multiplexer and further down-converted by a programmable frequency divider realized using cascaded D-type flip-flops, which use current-mode logic (CML) for speed. The outputs are fed off-chip using resistively-loaded differential buffers that are designed to satisfy the electrical requirements of the widely-used LVDS protocol. These include a differential impedance of 100 Ω , a differential amplitude of 250–450 mV, and a common-mode range of 0–2.4 V.

F. Design Summary

Table I summarizes the parameters of the major on-chip blocks, including those that are programmable via SPI. Each stage is independently programmed using its own address.

IV. SIMULATIONS AND MEASUREMENTS

The RF cochlea chip was fabricated in the UMC 65 nm CMOS process. The die measures 3.95 mm $\times$ 1.88 mm,

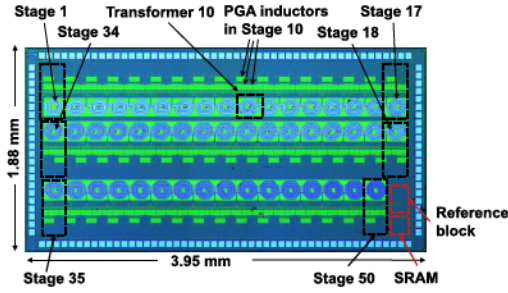
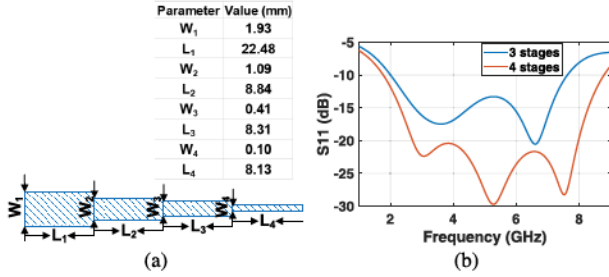


Fig. 13. Die photograph of the digitally-programmable RF cochlea chip.

Fig. 14. (a) Layout of a four-section 1-to-4 Chebyshev matching transformer; and (b) comparison of the simulated $|S_{11}|$ of three- and four-section designs.

as shown in Fig. 13. To eliminate package parasitics, dies were directly wire-bonded to test boards using a chip-on-board (COB) assembly process and then encapsulated for protection.

A. Test Board Design

To match the expected input impedance $Z_{in} = 200 \Omega$ of the cochlea chip to the 50Ω used by standard RF components and test equipment, a custom 1-to-4 impedance transformer was designed using microstrip lines on the test board. The design targets a Chebyshev response to maximize bandwidth at the expense of some passband ripple [36]. Layout of a four-section version is shown in Fig. 14(a), while Fig. 14(b) compares the reflection coefficients of three-section and four-section versions. Matching bandwidth increases with the number of sections: three and four sections provide $|S_{11}| < -10$ dB over 2.0-7.6 GHz (1:3.8) and 1.75-8.8 GHz (1:5.0), respectively. The latter was judged to be sufficient for this application.

The $N = 50$ LVDS outputs of the cochlea were interfaced to two high-speed mezzanine card (HSMC) connectors via 100Ω differential microstrip lines. This design allows the cochlea to be directly connected to a compatible FPGA development board. However, an additional interface board was designed and used during initial testing. This board, which also plugs into the HSMC connectors, contains a set of programmable 8-to-1 RF multiplexers and baluns that i) select 8 of the 50 differential LVDS outputs, and ii) convert them to single-ended form. The resulting experimental setup is shown in Figs. 15(a) and (b). Both the cochlea and the interface board are programmed by a microcontroller (Arduino Due).

B. Chip Characterization

1) *Input Impedance*: Fig. 16(a) shows the input reflection coefficient $|S_{11}|$ of the cochlea test board for different bias

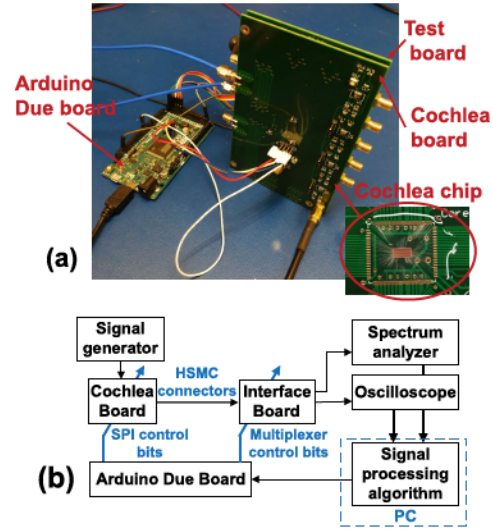
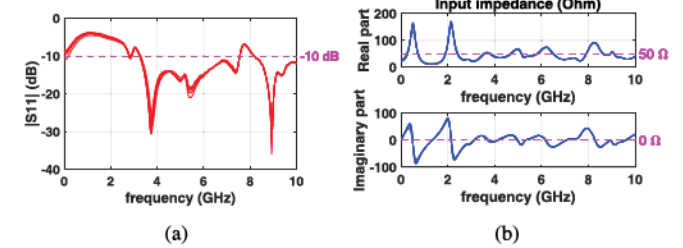


Fig. 15. Experimental setup: (a) photograph and (b) block diagram.

Fig. 16. Measured (a) $|S_{11}|$ parameter, and (b) input impedance of the cochlea test board at different bias current settings for the negative resistor R_1 .

current settings of the negative resistor R_1 ; the latter has a relatively small effect on $|S_{11}|$. The corresponding Z_{in} is shown in Fig. 16(b). Within the bandwidth of the matching transformer (1.75-8.8 GHz), Z_{in} is close to $Z_{in, cochlea}/4$, as expected (i.e., $\text{Re}(Z_{in}) \approx 50 \Omega$ and $\text{Im}(Z_{in}) \approx 0$).

2) *Frequency Encoding*: As explained in the previous section, the free-running frequencies of the divide-by-3 ILFDs in the frequency encoding section are set using two levels of tuning (coarse and fine). Fig. 17(a) shows the simulated free-running frequencies of the four ILFD banks, which serve as coarse tuning elements. In particular, banks “A”, “B”, “CD” and “EF” correspond to stages 1-8, 9-17, 18-34, and 35-50, respectively. Within each bank, each ILFD can be fine-tuned by its own 3-bit current DAC, as shown in the figure. Thus, the free-running and input frequencies are densely distributed over a broad range (approximately 0.266-4.15 GHz and 0.80-12.45 GHz, respectively). This range exceeds that of the cochlea, ensuring that signals near the best frequencies $\omega_c(n)$ of all cochlear stages can be detected.

Fig. 17(b) compares the average ILFD free-running frequency for bank “CD” versus bias current in three cases: simulated (both pre- and post-layout) and measured. The measurements are in good agreement with post-layout simulations.

ILFD locking range depends on the amplitude of the injected signal, as discussed earlier. The measured locking sensitivity curve of the divide-by-3 ILFD in stage 20 with a free-running frequency of 1.05 GHz is shown in Fig. 18. The circuit locks when the input power level at any particular

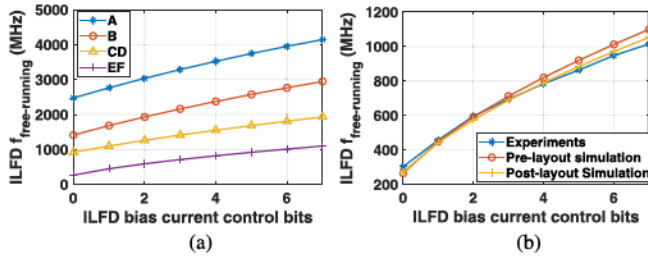


Fig. 17. (a) Simulated free-running frequency of different ILFD blocks versus bias current; (b) measured and simulated average free-running frequency of the “CD” ILFD block versus bias current.

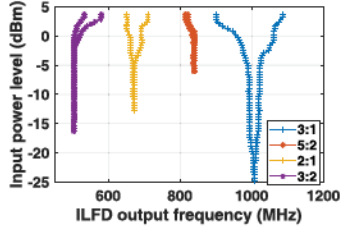


Fig. 18. Measured ILFD input sensitivity map for stage 20 (free-running frequency at 1.05 GHz), showing multiple Arnold tongues.

frequency exceeds the plotted value at that point. Besides the expected locking frequency ratio of 3:1, the ILFD also locks at other frequency ratios formed from simple whole numbers (e.g., 5:2, 2:1, and 3:2), but over narrower ranges. Similar behavior has been observed in other driven oscillator systems, and the parameter ranges within which locking occurs are known as Arnold tongues [37]. The measured locking range for the 3:1 tongue is approximately 10% (2.88–3.18 GHz) for a relatively high input power level of 0 dBm, and becomes even smaller as the power level decreases. This result highlights the key advantage of injection-locked oscillators for spectrum analysis, namely high frequency selectivity for weak signals.

Fig. 19(a) and (b) compares the measured output power spectrum and single-sided phase noise spectrum for an ILFD in the unlocked and locked states. The input to the cochlea was a low-noise continuous-wave (CW) signal at $3 \times 0.96 = 2.88$ GHz. Fig. 19(a) shows that the power spectrum becomes significantly narrower after locking, as expected. Also, the input was slightly offset from $3 \times f_0$, where f_0 is the free-running frequency, so locking results in a shift of the average output frequency. The expected reduction in phase noise after locking is confirmed in Fig. 19(b). In particular, the phase noise spectrum of the locked ILFD closely tracks that of the input signal (the two are <4 dB apart for offset frequencies up to ~ 1 MHz)¹, while that of the unlocked ILFD exhibits a flat region at low offset frequencies due to frequency drift during the measurement. Reduced phase noise is also accompanied by a large decrease in time-domain stability metrics such as

¹Since the input signal frequency is $3 \times$ higher than the locked output, the actual degradation in close-in phase noise after the ILFD detector is approximately $20 \log_{10}(3) + 4$ dB = 13.5 dB. Such degradation is likely caused by noise added to the input signal as it propagates through many cochlear stages before injection-locking a particular ILFD.

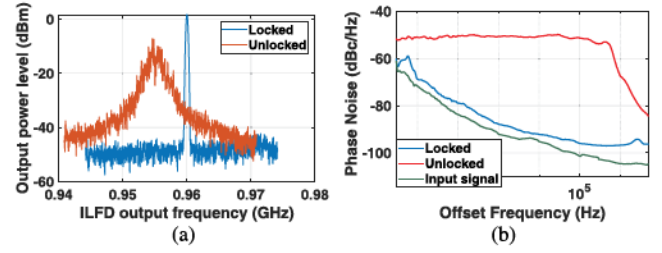


Fig. 19. Measured (a) output power spectrum and (b) phase noise spectrum (offset frequency from 3 Hz to 5 MHz) for an ILFD output in the unlocked and locked states. Phase noise spectra were averaged $N = 128$ times to reduce the displayed average noise level (DANL).

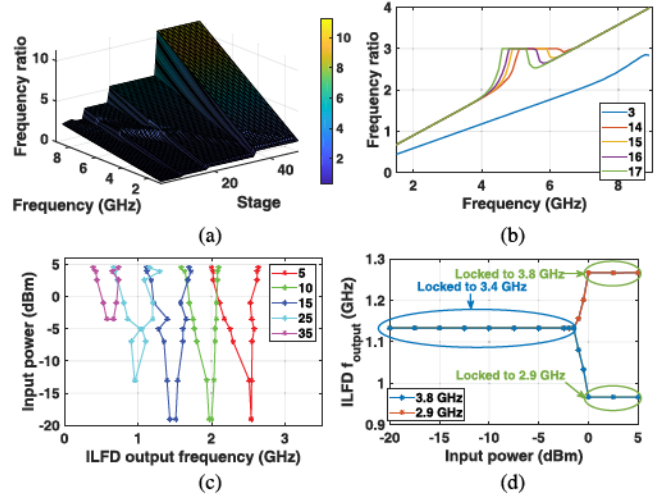


Fig. 20. (a) Simulated spatial frequency transfer functions of the cochlea to CW inputs (−4 dBm) at various frequencies; (b) measured spatial frequency transfer functions of the cochlea to CW inputs (0 dBm) at stages {3, 14–17}; (c) measured input sensitivity curves at stages {5, 10, 15, 25, 35}; (d) measured ILFD output frequency at stage 22 for a two-tone input consisting of a signal (3.4 GHz fixed at −10 dBm) and interferer (either 3.8 GHz or 2.9 GHz, varying from −20 to 5 dBm).

the Allan deviation $\sigma_A(\tau)$. The latter can be monitored using a frequency counter to quickly verify the ILFD’s locking state.

Fig. 20(a) shows the simulated spatial frequency transfer functions (i.e., ratio between ILFD input and output frequencies) of the cochlea as the frequency of a relatively strong CW input (−4 dBm) is swept from 1.0–9.0 GHz. Note that the free-running frequencies within each ILFD bank were set to nominally identical values during this simulation, resulting in the obvious steps visible between the “A”, “B”, “CD” and “EF” banks. In a practical scenario, these frequencies would be individually programmed to more closely match the exponentially-decreasing cochlear analysis frequency $\omega_c(n)$. Nevertheless, only a small set of cochlear stages lock at each input frequency, and these regions (visible as the “bumps” in the lower half of the figure) smoothly move towards the later stages (the apex) as the frequency decreases, thus verifying the expected logarithmic frequency-space mapping.

Fig. 20(b) shows the measured spatial frequency transfer functions at stages {3, 14–17} in more detail versus input frequency. Locking occurs when the transfer function (i.e., frequency ratio) = $X3$. As expected, stage 3 locks at a

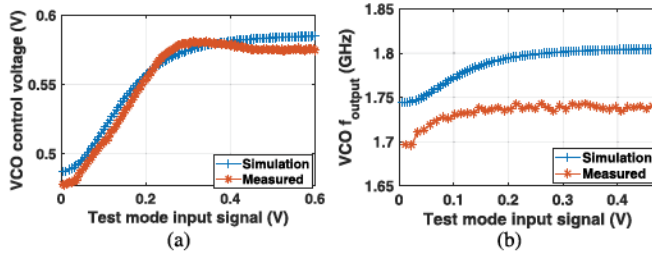


Fig. 21. Simulated and measured (a) ED output voltage, and (b) VCO output frequency as a function of test mode signal amplitude at 3.7 GHz.

high frequencies (>8.5 GHz), while stages 14-17 lock around 5 GHz. Within this set, the ILFD locking range also moves towards lower frequencies (to the left) as the stage number increases, as expected from cochlear frequency scaling.

Fig. 20(c) shows the measured input sensitivity curves over a broader set of stages, namely {5, 10, 15, 25, 35}. The Arnold tongues shift to lower frequencies as we move towards the apex (i.e., the stage number increases), again as expected. The minimum input power level required for locking also increases towards the apex due to signal attenuation in the cochlear transmission line. The associated losses are dominated by the series resistance of the transformer windings. The total amount of loss is ~ 15 dB, which is acceptable for many applications and comparable to our earlier work [9].

We used two-tone tests to study the interference-rejection capabilities of the frequency-encoded outputs. In these tests, the amplitude of a desirable input tone (signal) remained constant, while that of an undesired tone (interferer) was gradually increased. Fig. 20(d) shows the results of one such test, where the signal at 3.4 GHz remained fixed at -10 dBm, while a higher-frequency interferer (at 3.8 GHz) increased from $P_{int} = -20$ to 5 dBm. The ILFD at stage 22 (which is nominally tuned to 3.4 GHz) remains locked to the signal when $P_{int} < -1.5$ dBm, and finally locks to the interferer when $P_{int} > -1$ dBm. Thus, in this case the ILFD provides ~ 10 dB of interference rejection at a frequency offset of 400 MHz, in general agreement with the single-tone sensitivity curves shown in Fig. 20(c). Similar results were obtained when the test was repeated with lower-frequency interferers. For example, Fig. 20(d) shows that ~ 10 dB of rejection is available for an interferer at 2.9 GHz (i.e., at an offset of -500 MHz).

3) Amplitude Encoding: The characteristics of the ED and VCO in the amplitude encoding path were tested using a test mode input (enabled by a control bit). The amplitude of this signal (at 3.7 GHz) was swept, and the ED and VCO outputs (DC voltage and frequency, respectively) recorded as shown in Fig. 21. Both encoding curves are in good agreement with simulations, with an input-referred linear range of ~ 0.2 V. However, a small frequency offset of $\sim 2.9\%$ (50 MHz) is visible in the VCO output. Similar offsets (likely due to mismatch and additional layout capacitance) between the outputs of the cochlear stages were removed through a single-point calibration procedure. Specifically, the VCO frequency of each stage was recorded with no input signal present, stored in a look-up table, and subtracted from subsequent readings.

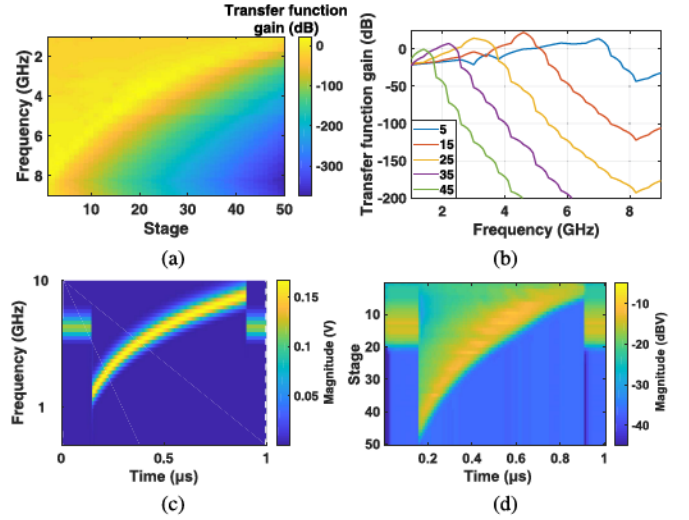


Fig. 22. Simulated small-signal spatial amplitude transfer functions of the cochlea to CW inputs (a) at various frequencies for all stages; and (b) at stages {5, 15, 25, 35, 45}. Simulated response to two simultaneously applied input frequencies (one input fixed at 4.32 GHz while the other increases from 1 GHz to 8.3 GHz): (c) time-frequency plot of the input, computed using a CWT; and (d) decoded amplitude outputs for all 50 cochlear stages.

Fig. 22(a) and (b) show the simulated small-signal spatial amplitude transfer functions of the cochlea to CW inputs at various frequencies from 1.0 GHz to 9.0 GHz. To obtain accurate results, the transformers were modeled using broadband 4-port S-parameter files extracted from Sonnet EM simulations. As expected, these functions have asymmetric band-pass shapes with cutoff locations that move logarithmically towards later stages (to the right of the figure) as the frequency decreases; the peak voltage gain is ~ 15 dB. Note that some ripple is visible in the cutoff responses: these are due to inter-stage reflections, which are not negligible due to the relatively high value of $Q_{line} = 0.08$ used in the design.

Fig. 22(c) shows the continuous wavelet transform (CWT) of a two-tone input with one input fixed at 4.32 GHz while the other sweeps from 1 GHz to 8.3 GHz linearly with time. Note that the CWT was chosen since the outputs of our cochlear model (in the linear regime) are well-modeled as constant- Q frequency bins, which resemble those generated by a CWT [38]. The cochlear amplitude outputs for all 50 stages generated by this input were decoded as functions of time and are plotted in Fig. 22(d). As expected, the cochlear time-frequency response is similar to a CWT of the input signal (see Fig. 22(c)). This result shows that the cochlear model indeed behaves as a real-time spectrum analyzer.

The measured output amplitudes for stages 18-24 (after calibration and decoding) for single-tone inputs at different frequencies are shown in Fig. 23(a). For an input at 3.8 GHz, the response peaks at stage 19 in agreement with simulations; the peak amplitude is ~ 3.2 dB larger than for the adjacent stages. Distinct peaks are observed for multi-tone inputs with well-separated components, thus allowing simple inference of these frequencies from the measured spatial output profiles. As an example, Fig. 23(b) shows measured profiles for inputs at 3.4 and 3.8 GHz (peaks at stages 19 and 22, respectively) as the amplitude of the high-frequency tone is varied.

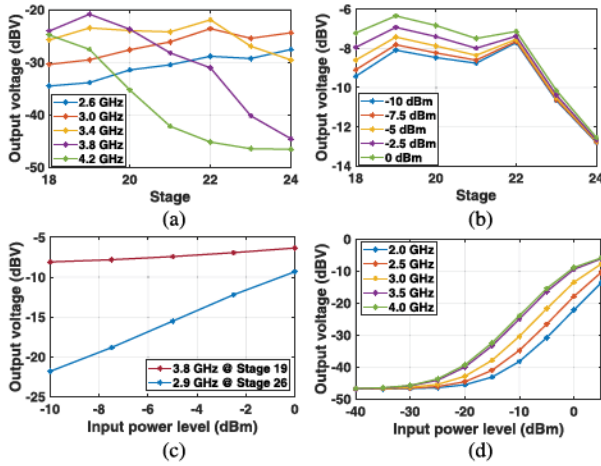


Fig. 23. Decoded output amplitudes at stages 18–24 for (a) single-tone inputs at different frequencies, all at -10 dBm; and (b) two-tone inputs (3.8 GHz varying from -10 to 0 dBm, and 3.4 GHz fixed at 0 dBm); (c) Comparison of decoded amplitudes for low- and high-side blocking: i) stage 18 (3.8 GHz signal, 3.4 GHz blocker), and ii) stage 26 (2.9 GHz signal and 3.4 GHz blocker). In both cases, the input amplitude varies from -10 to 0 dBm, while the blocker is fixed at 0 dBm. (d) Measured single-tone compression curves for stage 18 at different frequencies.

The experiments shown in Figs. 23(a)–(b) were performed at relatively low input amplitudes, resulting in approximately linear encoding. Next we consider nonlinear behavior at larger amplitudes. The low-pass nature of the cochlear transfer functions results in asymmetric gain compression due to strong inputs (blockers). In particular, low-frequency blockers propagate further along the cochlea and are thus more effective at causing gain compression than high-frequency ones. For example, Fig. 23(c) compares the decoded amplitudes from a single stage in two cases: i) low-side blocking of stage 19 by a 3.4 GHz tone (fixed at 0 dBm) as the desired signal at 3.8 GHz increases from -10 to 0 dBm; and ii) high-side blocking of stage 26 when the same 3.4 GHz tone blocks a desired signal at 2.9 GHz that increases from -10 to 0 dBm. Clearly, low-side blocking is much more effective (as shown by the nearly constant output amplitude versus input power level) than high-side blocking, as expected. Note that the frequency-encoded outputs have significantly better immunity to both low- and high-side blockers, as described in the previous section: this is due to the nonlinear locking dynamics of the ILFDs.

Fig. 23(d) shows how the gain of the cochlear transfer functions decreases with increasing input amplitude. These compression curves were taken by observing the single-tone response of stage 18 at various frequencies, including its best frequency of $f_{max} = 4.0$ GHz. The response at f_{max} , being larger, compresses at smaller input power levels than at other frequencies. This behavior is in agreement with simulations, and also qualitatively similar to that of the biological cochlea.

The analog outputs of each cochlear stage are quantized prior to off-chip transmission. Fig. 9(c) shows the basic architecture of the VCO-based quantizer used for this purpose. In one example, an amplitude-modulated signal with $f_{carrier} = 4.32$ GHz and $f_{signal} = 0.2$ MHz was applied at the input, and the measured output signal sampled at a frequency of f_{clk} . The measured ENOB with the input power fixed at

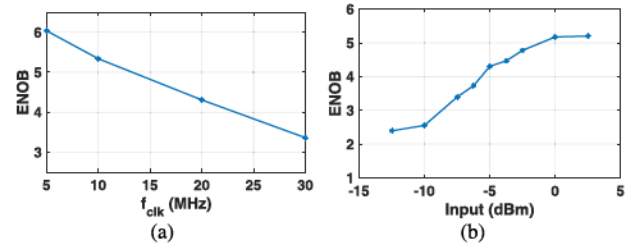


Fig. 24. Estimated ENOB for an amplitude-modulated input ($f_{carrier} = 4.32$ GHz and $f_{signal} = 0.2$ MHz) versus (a) sampling frequency f_{clk} (input fixed at -5.0 dBm), and (b) input power (f_{clk} fixed at 20 MHz).

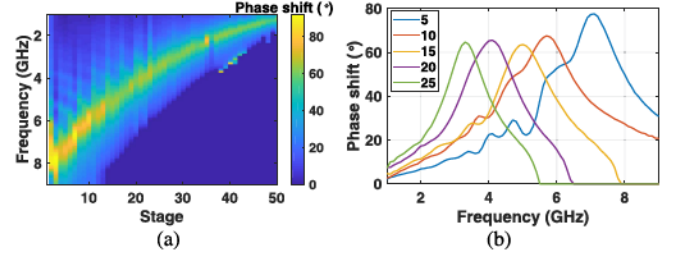


Fig. 25. Simulated spatial phase shift of the cochlea to CW inputs (a) at various frequencies; and (b) at stages {5, 10, 15, 20, 25}.

-5.0 dBm decreases as f_{clk} increases because of increased in-band quantization noise, as shown in Fig. 24(a). Also, the ENOB with f_{clk} fixed at 20 MHz increases as the input power increases, and reaches a maximum value of ≈ 5.2 bits as shown in Fig. 24(b). This level of precision is sufficient for digitizing typical wireless signals, which are received with relatively low SNR [39].

4) *Phase Encoding*: The spatial phase shift $\Delta\phi$ between stages for the cochlear model with CW inputs at various frequencies is shown in Fig. 25(a). The values of $\Delta\phi$ peak at nearly the same positions as in the amplitude response. However, unlike the latter, the phase responses have band-pass characteristics with sharp roll-offs on both sides of the peak; this results in higher frequency selectivity. The difference is visible if we compare Fig. 25(b), which plots $\Delta\phi$ for CW inputs at stages {5, 10, 15, 20, 25}, with Fig. 22(b), which plots the small-signal amplitude responses for the same stages.

The fact that $\Delta\phi < 80^\circ$ is used as a verification step during phase decoding. First, the VCO output frequencies are decoded into phase shifts using polynomial fits to the encoding curves. Next, decoded data with $\Delta\phi > 80^\circ$ is detected and ignored as invalid, and only valid data is converted to output phase shift. Figs. 26(a) and (b) shows simulated and measured phase shifts as a function of stage number (i.e., spatial position) for several input frequencies. The measured data is close to the simulation for stages where the ILFDs are locked, i.e., near the peaks of the amplitude transfer functions. For other stages, the measured value of $\Delta\phi$ is close to zero since the ILFDs are unlocked, i.e., generate random phase shifts that are averaged out by the LPF. This behavior is not observed in Fig. 26(a) due to a simulation artifact, namely that all ILFDs remain coherent in a noiseless transient simulation.

5) *Modulated Inputs*: ILFDs in the locked state can track input signals with different modulation schemes. Fig. 27(a) (left axis) shows the measured instantaneous frequency of

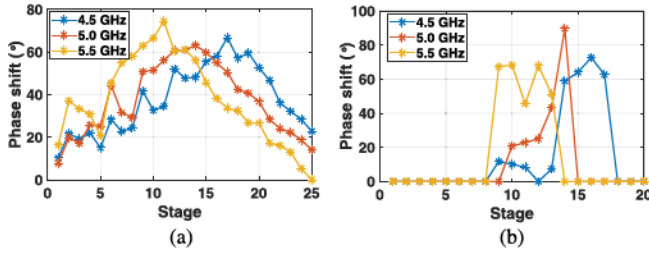


Fig. 26. (a) Simulated and (b) measured phase shift of the cochlea chip to inputs at 4.5, 5.0, and 5.5 GHz.

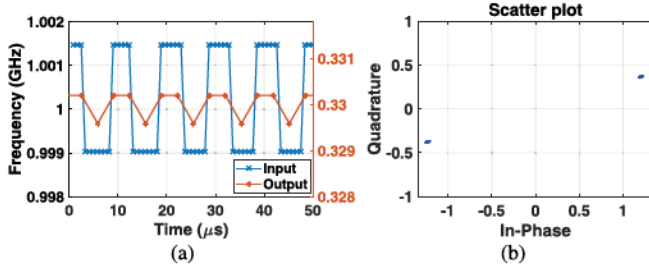


Fig. 27. Measured (a) FSK2 input and ILFD output waveforms from stage 50 (test mode) in the locked state; and (b) constellation diagram of the output.

a binary frequency-shift keying (FSK2) input signal with a carrier frequency of 1.0 GHz, a data rate of 0.2 Mbps, a frequency deviation of 2.5 MHz, and a power level of 0 dBm. Frequencies were estimated from a time-domain waveform using short-term Fourier transforms over moving windows. Fig. 27(a) (right axis) shows that the locked ILFD output tracks the input frequency deviations, as expected. The measured constellation of the output waveform is shown in Fig. 27(b); the error vector magnitude (EVM) is 0.85%.

C. Cognitive Receiver

The RF cochlea chip-based front-end discussed above provides us with real-time wideband time-frequency data. However, addressing RF spectral efficiency and security challenges requires a CR to combine real-time spectrum awareness with next-generation dynamic spectral access (DSA) algorithms. Real-time machine learning (ML) is a promising approach for combining spectrum awareness with DSA, as evidenced by the recent interest in deep learning for RF systems [40]–[42].

Here we describe a single-antenna CR receiver based on combining the RF cochlea chip with a deep learning algorithm, as shown in Fig. 28. It includes i) a broadband RF front-end (Vivaldi antenna and low-noise amplifier [LNA]); ii) the RF cochlea, which provides spectral awareness; iii) a trained deep belief network (DBN) [43] for modulation recognition (MR) [44]; and iv) a feedback loop for programming the cochlea based on the DBN outputs and user-specified goals (e.g., selecting signals with particular frequency bands and modulation types, while rejecting others). Such closed-loop operation enables real-time decision-making and learning, which is valuable for CR networks [40], [41].

During the tests, MATLAB was used to generate the frequency-time spectra of wideband communication signals based on one of several common digital modulation schemes (e.g., FSK2, FSK4, QAM2, QAM4, etc.). The resulting

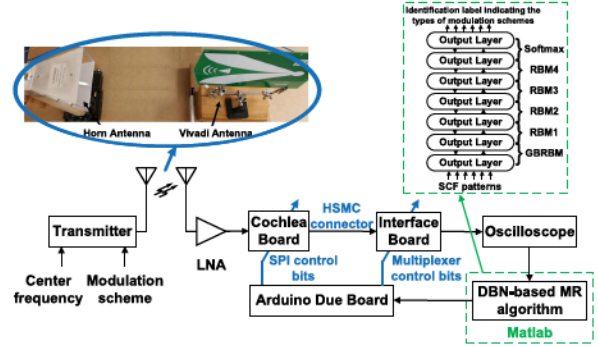


Fig. 28. Block diagram of an experimental prototype of a single-channel CR receiver based on the RF cochlea chip.

complex waveforms were uploaded to a dual-channel baseband function generator. The two output channels were then fed into the I - Q modulator of a RF vector signal generator. The modulated signal was then transmitted by a broadband (2-18 GHz) horn antenna and received by a planar Vivaldi antenna (0.9-12 GHz) placed ~ 0.5 m away. The received signals were amplified by the LNA and fed into the cochlea chip. The outputs of selected cochlear channels were digitized and fed into a DBN-based MR algorithm, also implemented in MATLAB as in our prior work [45]. The DBN was trained for up to 100 iterations using a gradient descent optimizer with a learning rate of 10^{-3} . The training data set consisted of 300 spectral correlation function (SCF) patterns for each modulation scheme [46]. The detection accuracy after training the network with these patterns was $> 80\%$ for all schemes.

The MR results were then used to adaptively isolate undesired wireless signals (blockers). For example, here we simultaneously transmit two modulated signals with the same power level (FSK2 at 4.32 GHz and QAM4 at 3.5 GHz, combined by a broadband power combiner before transmission) as shown in Fig. 29(a). Selected parallel outputs of the cochlea chip are kept under observation. At low input power when the ILFDs are unlocked, the modified Allen deviation $\sigma_A(\tau)$ has an average value of $\sim 4 \times 10^{-2}$; this decreases $\sim 10\times$ in the locked state. Individual channels are assumed to lock once $\sigma_A(\tau)$ falls below a preset threshold of $\sim 1 \times 10^{-3}$, which is $4\times$ lower than the average unlocked value (as shown in Fig. 29(b)). The trained DBN then classifies the results and optimizes the corresponding PGAs within the cochlea chip to maximize signal gain when the desired scheme is detected.

Fig. 29(c) shows the FSK2 signal after frequency encoding by stage 18 (including frequency division by a factor of $4.32/0.09 = 48$), and Fig. 29(d) shows the QAM4 signal after amplitude encoding by stage 22. These test results show that cochlear filtering enables the detectors to focus on to signals near the best frequency of each cochlear stage, effectively suppressing other signals (i.e., blockers).

D. Performance Summary

Table II summarizes the performance of this design and compares it with prior work. While it is based on a similar bidirectional cochlear model as [9], the new chip adds i) digital programmability, and ii) parallel digital outputs with

TABLE II
PERFORMANCE SUMMARY COMPARED WITH PRIOR WORK

Reference	CMOS technology	Frequency range	Peak voltage (dB)	Power (mW)	Dynamic range (dB)	Area (mm ²)	Digital programmability	Parallel output
This work	65 nm	1.0-8.3 GHz	12	246*	62	8	Yes	Yes
JSSC 2016/51.11 [18]	180 nm	0.008-20 kHz	/	0.55	60	50.4	Yes	Yes
TBCAS 2013/8.4 [17]	350 nm	0.050-50 kHz	/	14	52	13.74	Yes	No
JSSC 2009/44.6 [9] (Bidirectional)	130 nm	1.2-8 GHz	10	306	60	4.5	No	No
JSSC 2009/44.6 [9] (Unidirectional)	130 nm	0.6-6 GHz	35	286	35	4.5	No	No
TCAS-I 2008/55.4 [19]	Discrete	20-90 MHz	/	/	/	3.1×10^4	No	Yes

* The power consumption of the output LVDS buffer is not included.

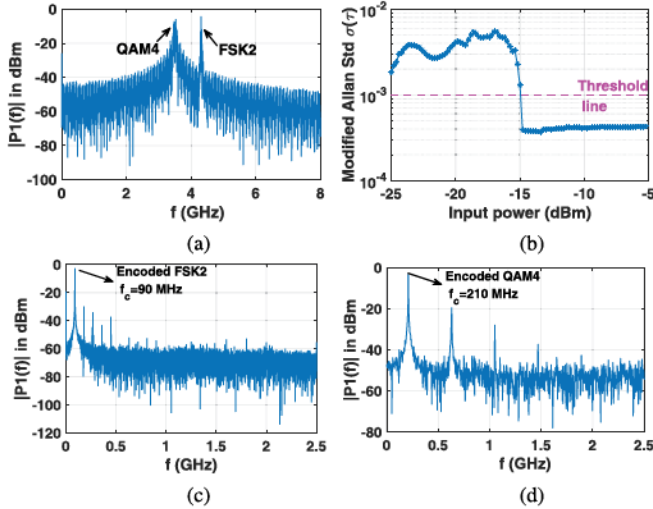


Fig. 29. (a) Measured power spectrum of an input signal consisting of two modulation types (FSK2 and QAM4); (b) modified Allan deviation of a single stage output (for averaging time $\tau = 6.4$ nsec) as a function of input power level; (c) frequency-encoded output spectrum at stage 18; and (d) amplitude-encoded output spectrum at stage 22.

amplitude, phase shift, or frequency encoding. Thus, it has much greater flexibility and $> 50\times$ higher output information rate compared to [9]: the total data rate for a sampling frequency of $f_{clk} = 20$ MHz is ~ 1 GS/s at 5-6 bits of precision. The chip also has higher dynamic range and lower power consumption (normalized to bandwidth) than recent audio-frequency silicon cochlea designs [17], [18]; this is largely due to its reliance on passive elements (transformers, capacitors, and resistors) instead of active circuits for cochlear pre-filtering. Finally, it has much higher bandwidth and smaller layout area than discrete (board-level) RF cochlea designs [19].

It is also instructive to compare the power consumption of the RF cochlea chip with the two conventional CR receiver front-end architectures shown in Figs. 1(a)-(b). Firstly, we would need multiple parallel down-converting channels in order for the heterodyne-based CR shown in Fig. 1(a) to provide the same bandwidth as our design. For example, a typical Ettus Research USRP X310 SDR has a real-time bandwidth of 160 MHz. We would need to run $7.3/0.16 = 45$ of these SDRs in parallel to monitor the entire 1.0-8.3 GHz analysis bandwidth of the RF cochlea, and the resulting power consumption would greatly exceed the 246 mW of our chip.

Secondly, in order for the direct-digitization CR architecture shown in Fig. 1(b) to match the 8.3 GHz upper frequency limit

of the cochlea, the system would have to sample at > 16 GS/s. Based on Murmann's comprehensive ADC survey [47], the state of the art Walden FOM for Nyquist-rate ADCs at 16 GS/sec is ≈ 30 fJ/bit. Assuming a similar ENOB = 6 bits as the cochlea, the expected ADC power consumption is ≈ 30 mW. The first operation performed on the digitized outputs is typically a low-resolution fast Fourier transform (FFT) to rapidly generate a coarse estimate of the spectrum. We estimated the power required to run such an FFT on a state of the art FPGA (Xilinx ZCU111 RFSoc, fabricated in 16 nm CMOS). In particular, we synthesized a 8-bit precise 32-point FFT running at 500 MHz, which is sufficient for monitoring a bandwidth of 8 GHz if its parallel inputs are generated by a polyphase filter bank. This initial DSP spectral estimation step was estimated to consume 3.55 W, which is already $10\times$ larger than that required by the RF cochlea to generate its analogous CWT-like spectral estimates. Thus, the proposed cochlea-based architecture (as shown in Fig. 1(c)) greatly reduces power consumption at the system level than these two conventional CR architectures by using analog pre-processing to extract meaningful features prior to digitization.

V. CONCLUSION

We have described a biologically-inspired broadband RF signal analyzer in 65 nm CMOS based on a transmission-line active cochlear model. This "RF cochlea" chip efficiently encodes amplitude, phase gradient, and frequency information from the cochlea using oscillators and ILFDs. This data can be used by ML models to extract higher-level information about the input signal, such as modulation types, that are valuable for CR applications. Moreover, the classification results can be fed back to the digitally-programmable RF cochlea chip to optimize the quality of the classification, adaptively suppress unwanted signals (e.g., blockers), and autonomously carry out other goal-directed tasks. A single-antenna CR prototype based on these principles was implemented and tested. Experimental results show i) $> 80\%$ over-the-air MR accuracy for several common schemes using a DBN classifier; and ii) autonomous self-optimization of the RF cochlea. Future work will focus on multi-antenna versions of such CR receivers.

REFERENCES

- [1] J. Mitola and G. Q. Maguire, "Cognitive radio: Making software radios more personal," *IEEE Pers. Commun.*, vol. 6, no. 4, pp. 13-18, Aug. 1999.

- [2] S. Haykin, "Cognitive radio: Brain-empowered wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 23, no. 2, pp. 201–220, Feb. 2005.
- [3] J. Park *et al.*, "A fully integrated UHF-band CMOS receiver with multi-resolution spectrum sensing (MRSS) functionality for IEEE 802.22 cognitive radio applications," *IEEE J. Solid-State Circuits*, vol. 44, no. 1, pp. 258–268, Jan. 2009.
- [4] T.-H. Yu, C.-H. Yang, D. Cabric, and D. Markovic, "A 7.4-mW 200-MS/s wideband spectrum sensing digital baseband processor for cognitive radios," *IEEE J. Solid-State Circuits*, vol. 47, no. 9, pp. 2235–2245, Sep. 2012.
- [5] N.-S. Kim and J. M. Rabaey, "A high data-rate energy-efficient triple-channel UWB-based cognitive radio," *IEEE J. Solid-State Circuits*, vol. 51, no. 4, pp. 809–820, Apr. 2016.
- [6] K. G. Smitha and A. P. Vinod, "A multi-resolution digital filter bank for spectrum sensing in military radio receivers," in *Proc. Int. Conf. Signal Process. Commun. (SPCOM)*, Jul. 2010, pp. 1–5.
- [7] Y. Hur *et al.*, "A wideband analog multi-resolution spectrum sensing (MRSS) technique for cognitive radio (CR) systems," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 2006, p. 4.
- [8] R. Sarpeshkar, *Ultra Low Power Bioelectronics: Fundamentals, Biomedical Applications, and Bio-Inspired Systems*. Cambridge, U.K.: Cambridge Univ. Press, 2010.
- [9] S. Mandal, S. M. Zhak, and R. Sarpeshkar, "A bio-inspired active radio-frequency silicon cochlea," *IEEE J. Solid-State Circuits*, vol. 44, no. 6, pp. 1814–1828, Jun. 2009.
- [10] G. Zweig, "Finding the impedance of the organ of corti," *J. Acoust. Soc. Amer.*, vol. 89, no. 3, pp. 1229–1254, Mar. 1991.
- [11] L. Watts, "Cochlear mechanics: Analysis and analog VLSI," Ph.D. dissertation, California Inst. Technol., Pasadena, CA, USA, 1992.
- [12] Y. Wang and S. Mandal, "Cochlear signal analysis for broadband spectrum sensing in cognitive radio networks," in *Proc. Cognit. Commun. Aerosp. Appl. Workshop (CCAA)*, Jun. 2017, pp. 1–4.
- [13] R. F. Lyon and C. Mead, "An analog electronic cochlea," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. 36, no. 7, pp. 1119–1134, Jul. 1988.
- [14] R. Sarpeshkar, R. F. Lyon, and C. Mead, "A low-power wide-dynamic-range analog VLSI cochlea," in *Neuromorphic Systems Engineering*. Boston, MA, USA: Springer, 1998, pp. 49–103.
- [15] R. Sarpeshkar, R. F. Lyon, and C. Mead, "A low-power wide-dynamic-range analog VLSI cochlea," *Anal. Integr. Circuits Signal Process.*, vol. 16, no. 3, pp. 245–274, 1998.
- [16] 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, Reg. Papers*, vol. 54, no. 1, pp. 48–59, Jan. 2007.
- [17] S.-C. Liu, A. van Schaik, B. A. Minch, and T. Delbruck, "Asynchronous binaural spatial audition sensor with $2 \times 64 \times 4$ channel output," *IEEE Trans. Biomed. Circuits Syst.*, vol. 8, no. 4, pp. 453–464, Aug. 2014.
- [18] M. Yang *et al.*, "A 0.5 V 55 μ W 64×2 channel binaural silicon cochlea for event-driven stereo-audio sensing," *IEEE J. Solid-State Circuits*, vol. 51, no. 11, pp. 2554–2569, Nov. 2016.
- [19] C. J. Galbraith, R. D. White, L. Cheng, K. Grosh, and G. M. Rebeiz, "Cochlea-based RF channelizing filters," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 55, no. 4, pp. 969–979, May 2008.
- [20] C. J. Galbraith and G. M. Rebeiz, "Higher order cochlea-like channelizing filters," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 7, pp. 1675–1683, Jul. 2008.
- [21] S. Mandal and R. Sarpeshkar, "A bio-inspired cochlear heterodyning architecture for an RF fovea," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 58, no. 7, pp. 1647–1660, Jul. 2011.
- [22] S. Mandal, S. M. Zhak, and R. Sarpeshkar, "Architectures for universal or software radio," U.S. Patent 8 121 223, Feb. 21, 2012.
- [23] S. Mandal, "Collective analog bioelectronic computation," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, May 2009.
- [24] S. Puria and J. B. Allen, "A parametric study of cochlear input impedance," *J. Acoust. Soc. Amer.*, vol. 89, no. 1, pp. 287–309, Jan. 1991.
- [25] A. G. Katsiamis, E. M. Drakakis, and R. F. Lyon, "A biomimetic, 4.5 μ W, 120+ dB, log-domain cochlea channel with AGC," *IEEE J. Solid-State Circuits*, vol. 44, no. 3, pp. 1006–1022, Mar. 2009.
- [26] W. Bialek and F. Rieke, "Reliability and information transmission in spiking neurons," *Trends Neurosci.*, vol. 15, no. 11, pp. 428–434, Nov. 1992.
- [27] S. Shamma, "On the role of space and time in auditory processing," *Trends Cognit. Sci.*, vol. 5, no. 8, pp. 340–348, Aug. 2001.
- [28] S. S. Mohan, "The design, modeling and optimization of on-chip inductor and transformer circuits," Ph.D. dissertation, Dept. Elect. Eng., Stanford Univ., Stanford, CA, USA, Dec. 1999.
- [29] C. P. Yue and S. S. Wong, "Design strategy of on-chip inductors for highly integrated RF systems," in *Proc. Design Autom. Conf.*, 1999, pp. 982–987.
- [30] G. Jovanovic, M. Stojcev, and Z. Stamenkovic, "A CMOS voltage controlled ring oscillator with improved frequency stability," *Appl. Math., Informat. Mech.*, vol. 2, no. 1, pp. 1–9, 2011.
- [31] 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 Syst. I, Reg. Papers*, vol. 57, no. 1, pp. 18–30, Jan. 2010.
- [32] S. Verma, H. R. Rategh, and T. H. Lee, "A unified model for injection-locked frequency dividers," *IEEE J. Solid-State Circuits*, vol. 38, no. 6, pp. 1015–1027, Jun. 2003.
- [33] J.-C. Chien and L.-H. Lu, "Analysis and design of wideband injection-locked ring oscillators with multiple-input injection," *IEEE J. Solid-State Circuits*, vol. 42, no. 9, pp. 1906–1915, Sep. 2007.
- [34] 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 *Proc. IEEE Custom Integr. Circuits Conf.*, Sep. 2009, pp. 259–262.
- [35] B. Razavi, "A study of injection locking and pulling in oscillators," *IEEE J. Solid-State Circuits*, vol. 39, no. 9, pp. 1415–1424, Sep. 2004.
- [36] D. M. Pozar, *Microwave Engineering*. Hoboken, NJ, USA: Wiley, 2009.
- [37] L. Glass, "Cardiac arrhythmias and circle maps—A classical problem," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 1, no. 1, pp. 13–19, Jul. 1991.
- [38] J. Yao and Y.-T. Zhang, "The application of bionic wavelet transform to speech signal processing in cochlear implants using neural network simulations," *IEEE Trans. Biomed. Eng.*, vol. 49, no. 11, pp. 1299–1309, Nov. 2002.
- [39] J. Singh, O. Dabeer, and U. Madhow, "On the limits of communication with low-precision analog-to-digital conversion at the receiver," *IEEE Trans. Commun.*, vol. 57, no. 12, pp. 3629–3639, Dec. 2009.
- [40] 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. IEEE*, vol. 102, no. 3, pp. 307–320, Mar. 2014.
- [41] T. J. O'Shea, L. Pemula, D. Batra, and T. C. Clancy, "Radio transformer networks: Attention models for learning to synchronize in wireless systems," in *Proc. 50th Asilomar Conf. Signals, Syst. Comput.*, Nov. 2016, pp. 662–666.
- [42] T. J. O'Shea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE J. Sel. Topics Signal Process.*, vol. 12, no. 1, pp. 168–179, Feb. 2018.
- [43] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, Jul. 2006.
- [44] G. J. Mendis, J. Wei-Kocsis, and A. Madanayake, "Deep learning based radio-signal identification with hardware design," *IEEE Trans. Aerosp. Electron. Syst.*, vol. 55, no. 5, pp. 2516–2531, Oct. 2019.
- [45] Y. Wang, X. Tang, G. J. Mendis, J. Wei-Kocsis, A. Madanayake, and S. Mandal, "AI-driven self-optimizing receivers for cognitive radio networks," in *Proc. IEEE Cognit. Commun. Aerosp. Appl. Workshop (CCAAW)*, Cleveland, OH, USA, Jun. 2019, pp. 1–5.
- [46] D. B. Cabric, "Cognitive radios: System design perspective," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., Univ. California, Berkeley, CA, USA, Dec. 2007.
- [47] B. Murmann. (2017). *ADC Performance Survey 1997-2017*. [Online]. Available: <http://web.stanford.edu/murmann/adcsurvey.html>