Ultra-Low-Power and Compact-Area Analog Audio Feature Extraction Based on Time-Mode Analog Filterbank Interpolation and Time-Mode Analog Rectification

Subhajit Ray, *Student Member, IEEE*, and Peter R. Kinget, *Fellow, IEEE*

Abstract— To address the power and area bottleneck imposed by the frontend feature extractor relative to the backend neural network in on-device keyword spotting (KWS), we propose two time-mode analog signal processing (ASP) circuit techniques showcased in an analog audio feature extractor chip that advances the state of the art in power- and area-efficiency. *Time-mode analog filterbank interpolation* uses digital XOR gates to double the number of outputs of an analog bandpass filterbank. *Time-mode analog rectification* uses a single digital XOR gate as an analog full-wave rectifier. The 65 nm low power (LP) CMOS chip uses only 80 nW and 0.53 mm^2 to extract from an input analog audio signal, an output digital auditory feature vector with 31 elements. This represents 18× and 3.3× improvements in power/feature and area/feature, as compared, respectively, to the most area- and power-efficient published analog audio feature extractor chips. All the while, competitive classification accuracy is maintained at >90% across ten keywords, as evaluated by feeding the chip's digital output directly into a small-footprint software backend classifier.

Index Terms— Always-on, analog feature extraction, analog signal processing (ASP), audio, audio recognition, feature extraction, keyword spotting (KWS), on-device, speech, speech recognition, time-domain circuits, time-mode circuits, ultra-low power.

I. INTRODUCTION

THE advent of ubiquitous computing calls for speech-controlled devices that are dispersed everywhere in the field. Such devices must be *always-on* in order to always be listening; must host their speech recognition *on-device* to avoid the energy, latency, and privacy issues that would arise from communication with cloud datacenters; and must be *battery-powered* to enable survivability in the field. With large-vocabulary continuous speech recognition as the ultimate goal, *keyword spotting (KWS)*—the task of recognizing a single spoken word—serves as a preliminary step. Therefore, the motivating application for this research is *always-on*, *on-device*, *battery-powered KWS*.
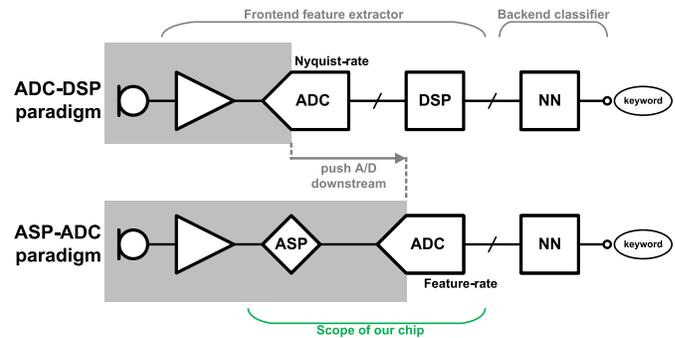
Fig. 1. Instead of the conventional ADC-DSP paradigm for KWS, we opt for the less conventional ASP-ADC paradigm, motivated by its relative power efficiency at the low SNRs (<50 dB) suitable for KWS. The scope of our chip is the analog-domain feature extraction plus feature-rate ADC, which we implement using the two proposed time-mode ASP circuit techniques.

For a decade-long lifetime on a typical coin cell battery (SR927 60 mAh, 1.55 V, 9.5 × 2.7 mm), the power budget is 1 μW. However, the lowest-power published KWS chip that is fully-integrated—receiving analog audio at its input, and producing a keyword label at its output—exceeds this budget by more than an order of magnitude, consuming 16.1 μW for 91%-accuracy 10-KWS [1]. Adopting the conventional analog-to-digital conversion plus digital signal processing (ADC-DSP) paradigm depicted at the top of Fig. 1, [1]'s frontend feature extractor—microphone pre-amplifier, ADC, digital-domain feature extraction—is the power bottleneck, consuming 2.6× more power than its backend classifier—a digital neural network. Therefore, research into power-efficient frontend feature extractors for KWS is justified.

More power-efficient than the conventional ADC-DSP paradigm is the analog signal processing plus ADC (ASP-ADC) paradigm, originating in the 1980s [2] for audio, and shown at the bottom of Fig. 1. ASP-ADC pushes the analog-digital boundary downstream to just before the classifier, enabling analog-domain feature extraction followed by not Nyquist-rate, but rather feature-rate [3] ADC. For two reasons, first principles suggest that ASP-ADC should be more power-efficient than ADC-DSP in the application of KWS audio feature extraction.

1) The feature-rate ADC should be lower power than the Nyquist-rate ADC, because the audio feature rate of 100 Hz [4] is slower than the audio Nyquist rate of 8 kHz.

2) The analog-domain feature extractor should be lower power than the digital-domain feature extractor because, more generally, ASP is more power-efficient than DSP at the signal-to-noise ratios (SNRs) of <50 dB [3], [5], [6], [7] suitable for KWS.

That ASP-ADC can be more power-efficient than ADC-DSP for KWS audio feature extraction has been demonstrated experimentally too: the most power-efficient ASP-ADC [8] is 12× more power-efficient than the most power-efficient ADC-DSP (ADC [9], DSP [10]). However, area comparisons have been largely ignored in the literature. In this example, ASP-ADC is 4.4× less area-efficient than ADC-DSP—a disadvantage that will only worsen with continued technology scaling. Therefore, *achieving both power- and area-efficiency in analog audio feature extraction, while maintaining classification accuracy through a digital backend classifier, remains a challenge.*

Toward addressing this challenge, we use a time-mode analog signal representation—where an analog value is encoded along the time axis in a digital pulsewidth that is continuous-valued—that enables digital gates to perform ASP, so as to simultaneously harness the low-SNR power-efficiency of analog, and technology scaling-driven area-efficiency of digital. Case in point, in such a time-mode analog signal representation, the digital XOR gate serves as an extremely efficient implementation of an analog subtraction merged with an analog absolute value [11], a property used in the two proposed time-mode ASP circuit techniques.

time-mode analog filterbank interpolation uses digital XOR gates to double the number of outputs an analog bandpass filterbank. This breaks the N filters $\rightarrow N$ features architectural limitation of previous analog audio feature extractors, improving it to $N \rightarrow 2N - 1$. *Time-mode analog rectification* uses a single digital XOR gate as an analog full-wave rectifier. This alleviates the power dissipation and area occupancy bottlenecks imposed by the rectifiers relative to the filters in previous analog audio feature extractors.

The remainder of the article, which elaborates on the analog audio feature extractor chip reported in [12], is organized as follows. Section II overviews the chip architecture, Section III develops time-mode analog filterbank interpolation and time-mode analog rectification, Section IV discusses their non-idealities, Section V details circuit design, Section VI reports experimental results based and discusses them, and finally, Section VII recapitulates the key points of the article.

II. CHIP ARCHITECTURE

A. Overview

Fig. 2 shows the chip architecture. Tracing the signal flow as follows.

1) The chip's input analog audio signal is analyzed into its bandpass components by a bandpass filterbank.

2) Each voltage-mode bandpass component is converted into a proportional time-mode analog signal by comparing it against a voltage-mode triangle wave using a continuous-time comparator (Section II-B).

3) Each green XOR gate computes an absolute-valued subtraction (Section II-C) between adjacent time-mode bandpass components, merging two functions: the creation of an "interpolated" time-mode bandpass component (Section III-A), and it full-wave rectification.

4) Each blue XOR gate computes an absolute value of a "native" time-mode bandpass component to full-wave rectify it (Section III-B).

5) The output duty cycle of each XOR gate is converted into a proportional spike rate by an integrate-and-fire (IAF), and a counter integrates this spike rate by counting the number of spikes over a window as a digital number, merging windowed integration with conversion to digital—a technique discussed in more detail in [13].

Finally, the chip output is a digital auditory feature vector, where each of its 31 elements is a 100 S/s number proportional to the amplitude integrated over a 20 ms window with 10 ms overlap in one of 31 sub-bands exponentially distributed across the speech band from 100 Hz to 4 kHz. Stacking feature vectors over time forms a digital auditory spectrogram.

B. Time-Mode Analog Signal Representation

Fig. 3 depicts the time-mode analog signal representation used in the chip architecture. In particular, it shows that a voltage-mode signal $V(t)$ is compared against a voltage-mode triangle wave $V_{\mathrm{tri}}(t)$ to produce a pulse-width modulated (PWM) waveform $V_{\mathrm{pwm}}(t)$ of the naturally-sampled asymmetric double edge flavor [14], which is interpreted to represent a fictitious discrete-time analog signal $T[k]$ whose values are equal to $V_{\mathrm{pwm}}(t)$'s pulsewidths. Specifically, a voltage of 0 maps to a $(1/2)T_{\mathrm{pwm}}$-width pulse, while voltages of $+(1/2)V_{\mathrm{tri}}^{\mathrm{pp}}$ and $-(1/2)V_{\mathrm{tri}}^{\mathrm{pp}}$ map to T_{pwm}-width and 0-width pulses, respectively.

In this way, it is $V_{\mathrm{pwm}}(t)$, or, equivalently, $T[k]$, that is the time-mode representation of $V(t)$. Observing that $T[k]$'s sampling frequency is inherited from $V_{\mathrm{tri}}(t)$ and equal to $V_{\mathrm{tri}}(t)$'s fundamental frequency f_{tri}, it follows that f_{tri} must be sufficiently faster than $V(t)$'s bandwidth by some factor in order for $T[k]$ to represent $V(t)$ with no aliasing. Had $V_{\mathrm{pwm}}(t)$ been of the uniformly-sampled variety, that factor would simply be 2; but because $V_{\mathrm{pwm}}(t)$ is naturally-sampled, and because for this case a derivation of the factor is lacking, we heuristically argue that it would be in the vicinity of the number two. This, therefore, suggests that given the speech bandwidth of 4 kHz, f_{tri} should be set to at least 8 kHz. However, recognizing that the target task is not perfect signal reconstruction of the input, but rather spotting keywords in the input, we surmise that some aliasing can be tolerated, and therefore f_{tri} relaxed; we empirically find that relaxing it from 8 to 4 kHz maintains greater-than-90% 10-word KWS classification accuracy, as reported in Section VI-B. Finally, although it is conceivable that the backend neural network could learn the signal distortion caused by slope nonlinearity
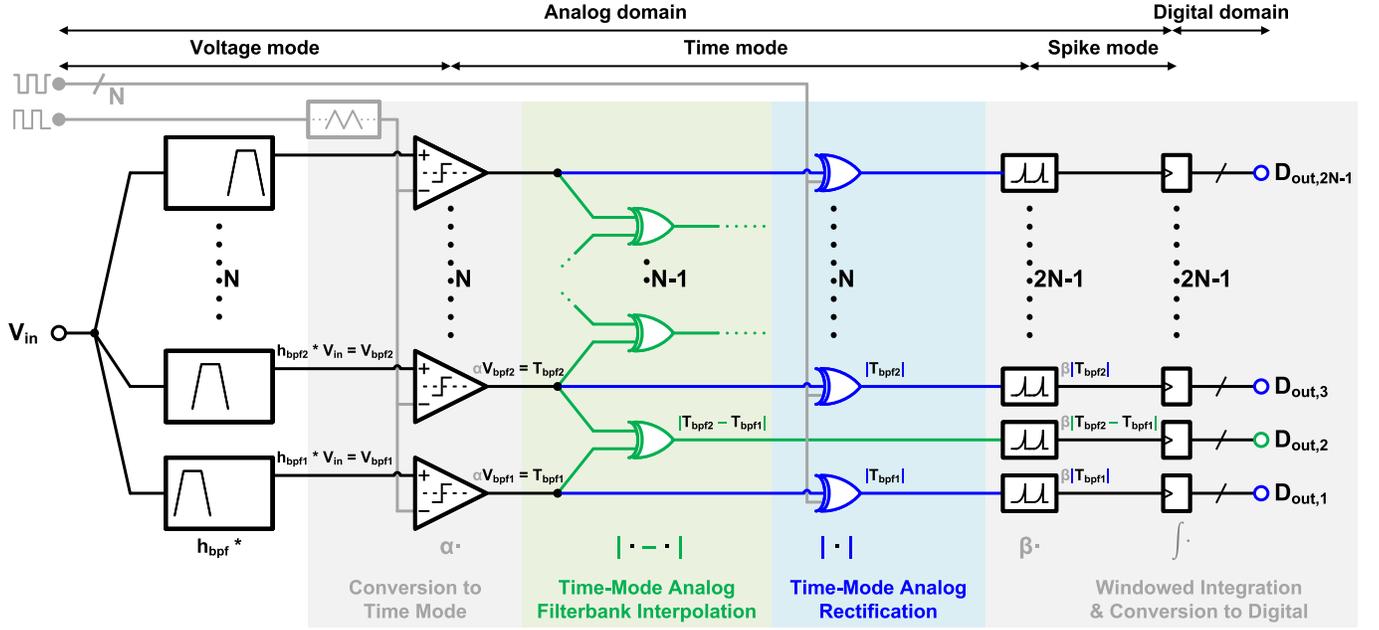
Fig. 2. Chip architecture, which combines the two proposed techniques of time-mode analog filterbank interpolation and time-mode analog rectification. The input is an analog audio signal and the output is a digital auditory feature vector.
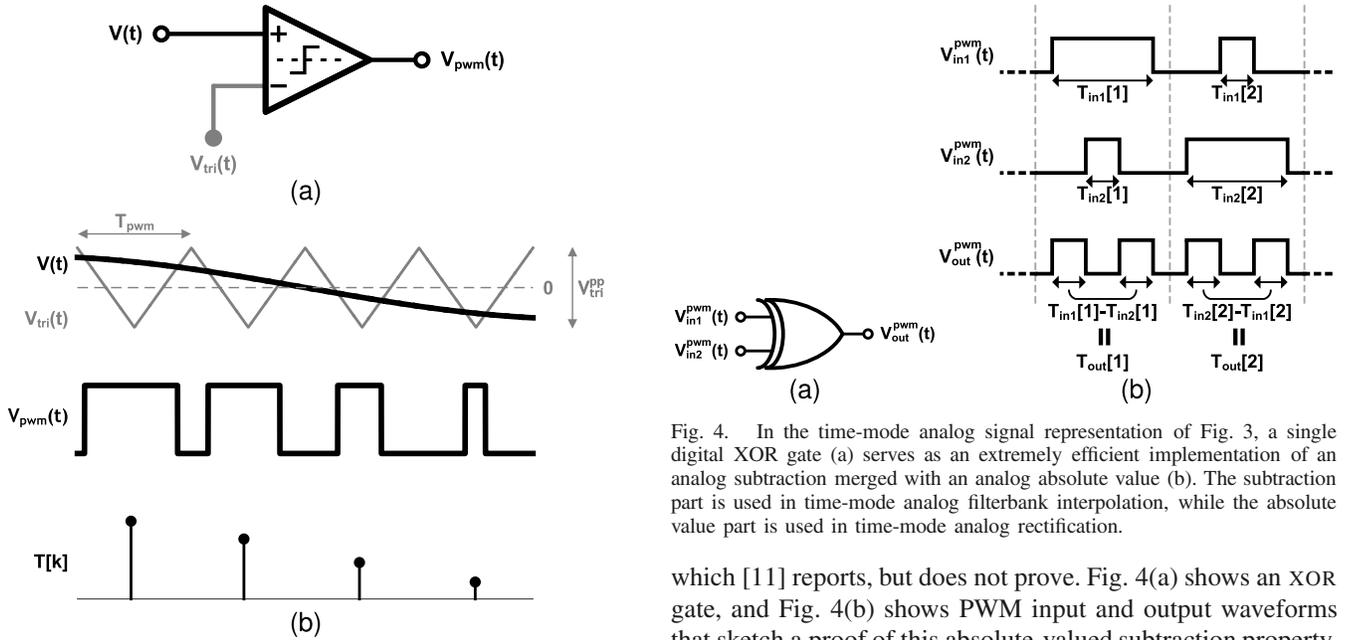

Fig. 3. Chip architecture uses a time-mode analog signal representation. (a) Voltage-mode signal $V(t)$ is converted into a PWM waveform $V_{pwm}(t)$ by a continuous-time comparator. (b) PWM waveform $V_{pwm}(t)$ is interpreted to represent a fictitious discrete-time analog signal $T[k]$ whose sample values are equal to the widths of the pulses and sample times are equal to the times of the centers of the pulses.

in $V_{tri}(t)$, we keep the absolute slope variation under a couple percent out of good practice.

C. XOR *Gate as Absolute-Valued Subtractor*

In the time-mode analog signal representation, the digital XOR gate functions as an analog absolute-valued subtractor,

Fig. 4. In the time-mode analog signal representation of Fig. 3, a single digital XOR gate (a) serves as an extremely efficient implementation of an analog subtraction merged with an analog absolute value (b). The subtraction part is used in time-mode analog filterbank interpolation, while the absolute value part is used in time-mode analog rectification.

which [11] reports, but does not prove. Fig. 4(a) shows an XOR gate, and Fig. 4(b) shows PWM input and output waveforms that sketch a proof of this absolute-valued subtraction property. Intuitively, when XOR gate's two inputs are fed with two different width pulses, it zeroes out the portion where the two pulses overlap, which amounts to a subtraction of pulsewidths; further, because the truth table of the XOR gate is symmetric between its inputs, when its inputs are interchanged, the output remains unchanged, implying the subtraction is absolute-valued.

III. TIME-MODE ANALOG TECHNIQUES

A. Time-Mode Analog Filterbank Interpolation

1) Advantages: The architecture of previously published state-of-the-art analog audio feature extractors [8], [15], [16],
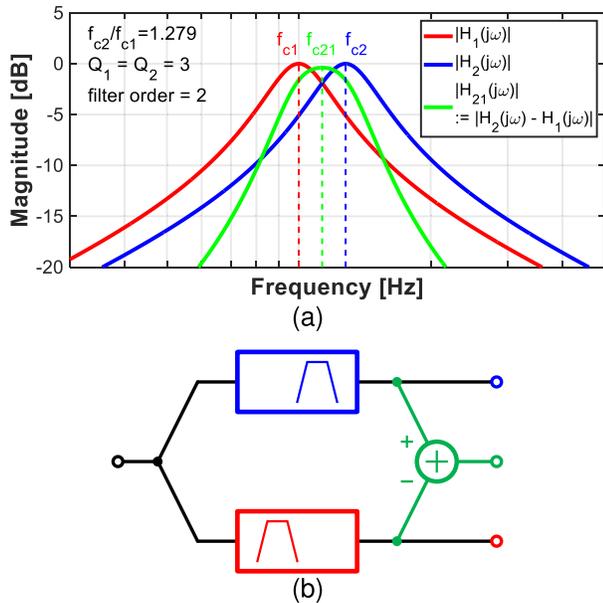
Fig. 5. Principle of time-mode analog filterbank interpolation. Subtracting two "native" bandpass transfer functions, $H_1(s)$ and $H_2(s)$: (a) creates an "interpolated" bandpass transfer function, $H_{21}(s)$ and (b) with the corresponding signal processing-level diagram. Fig. 2 shows that the subtractor in (b) is implemented using a XOR gate. In this way, digital XOR gates double the number of bandpass outputs of an analog bandpass filterbank.

[17], [18] is essentially the same as it was 40 years ago [2], requiring N bandpass filters to produce N bandpass features. Time-mode analog filterbank interpolation breaks this $N \rightarrow N$ limitation, improving it $N \rightarrow 2N - 1$, by simply inserting $N - 1$ digital XOR gates. For a target number of features, this reduces filterbank power and area by $2\times$ and $2\times$, and consequently total feature extractor power and area by $1.6\times$ and $1.7\times$, as compared to not using time-mode analog filterbank interpolation, but still using time-mode analog rectification (Section III-B).

2) Principle: Time-mode analog filterbank interpolation is the result of a signal processing insight and circuits insight. Fig. 5 depicts the signal processing insight, which is that the subtraction between two "native" bandpass transfer functions, $H_1(s)$ and $H_2(s)$, is another, "interpolated" bandpass transfer function $H_{21}(s) = H_2(s) - H_1(s)$ that has a magnitude response $|H_{21}(j\omega)|$ with a center frequency f_{c21} that is interpolated between those, f_{c1} and f_{c2}, of $|H_1(j\omega)|$ and $|H_2(j\omega)|$, and is given by $f_{c21} = (f_{c1} f_{c2})^{1/2}$. In this design, f_{c1}/f_{c2} is set to 1.279, as is required to exponentially distribute 16 center frequencies from 100 Hz to 4 kHz. Given $f_{c1}/f_{c2} = 1.279$, the native Q_1 and Q_2 are set to 3, as this results in an interpolated Q_{21} close in value at 2.7. Finally, second-order transfer functions are used for $H_1(s)$ and $H_2(s)$, as they have been shown to be sufficient for KWS [8].

Fig. 2 depicts the circuit's insight in its green shaded part, which is that the necessary subtraction in Fig. 5(b) can be implemented using the subtraction part of the XOR gate's absolute-valued subtraction.[1] Fig. 2 further shows that

[1]The absolute value part is used in time-mode analog rectification (Section III-B).

with the filters in Fig. 5(b) implemented as voltage-mode filters, subsequent conversion to time mode is necessary to interface the voltage-mode output of the filters with the time-mode input of the XOR gate. More generally, starting with N analog bandpass filters, time-mode analog filterbank interpolation uses $N - 1$ digital XOR gates to compute analog subtractions between adjacent bandpass filter outputs to create $N - 1$ interpolated bandpass features, which, together with the N native bandpass features from the N filters, make for a total of $2N - 1$ bandpass features.

B. Time-Mode Analog Rectification

1) Advantages: In general, analog audio feature extractors are composed of filters followed by rectifiers, and previous works suffer from the rectifiers consuming $3.3\times$–$19.8\times$ more power than the filters [13], [19]. Time-mode analog rectification replaces the conventional, complex, voltage/current-mode rectifiers used in both with a single digital XOR gate. This reduces rectifier power/area by $6.6\times/2.8\times$ (including the overhead of conversion to time-mode) as compared to [13], and consequently total feature extractor power and area in this work by $3.7\times$ and $1.9\times$, as compared to not using time-mode analog rectification (and, necessarily, not using time-mode analog filterbank interpolation).

2) Principle: Time-mode analog rectification is the result of the insight that the absolute value part of the XOR gate's absolute-valued subtraction can be used to compute an absolute value, i.e., full-wave rectification. This is done by feeding one input of the XOR gate with a 50% duty-cycle squarewave, which, because it is the time-mode representation of the number zero (Section II-B), causes the absolute-valued subtraction that the XOR gate would otherwise compute, to reduce to just an absolute value. Fig. 6 shows the XOR gate's transfer characteristic, from input pulsewidth to output pulsewidth, along with exemplary waveforms. Note that the frequency of the squarewave, $V_{\text{sqr}}(t)$, must be the same as the frequency of the PWM input to the XOR gate, $V_{\text{in}}^{\text{pwm}}(t)$, which itself is inherited from and equal to the triangle wave frequency, f_{tri}, which is 4 kHz.

Fig. 2 shows in the blue shaded part that, as with time-mode analog filterbank interpolation, the XOR gate must be preceded by conversion to time mode using a continuous-time comparator to interface its time-mode input with the filter's voltage-mode output.

IV. NON-IDEALITIES

In principle, each and every non-ideality of each and every circuit block contributes to the degradation of the quality of the chip's digital output features. However, at the nA current levels and kHz input signal bandwidth made possible and necessitated by the target KWS application, we empirically observe that the dominant non-idealities are those of the continuous-time comparators. In contrast, an example of a negligible circuit block non-ideality is the delay of the XOR gate. This is because the kHz input signal bandwidth allows the PWM period to be as long as on the order of 100 μs, whereas the delay of an XOR gate is as short as on the order of 10 ns (at the reduced 0.4 V digital supply the chip uses).
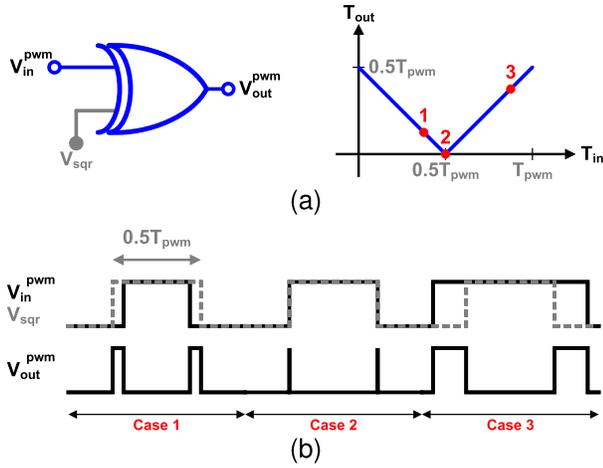
Fig. 6. Principle of time-mode analog rectification. A single digital XOR gate computes an analog absolute value, i.e., analog full-wave rectification, in that: (a) its input-pulsewidth-to-output-pulsewidth transfer characteristic is absolute valued-shaped and (b) as illustrated by exemplary waveforms.
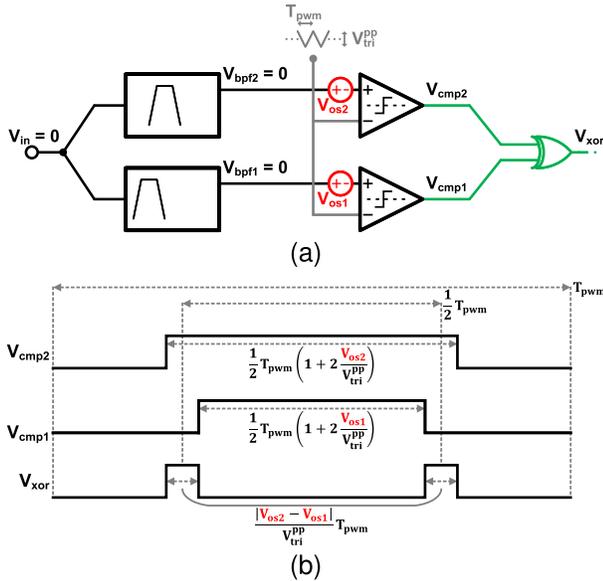

Fig. 7. Dominant non-ideality in time-mode analog filterbank interpolation is comparator offset (a), which causes the XOR gate to output nonzero pulsewidths (b) despite zero chip input. This limits the stopband rejection of the interpolated features. The rejection can be improved by realizing each comparator with many sub-comparators, then selecting the pair of sub-comparators that minimizes the XOR gate output duty cycle. Note that the effect of offset is exaggerated for illustration purposes.

A. Time-Mode Analog Filterbank Interpolation

The dominant non-ideality in time-mode analog filterbank interpolation is comparator offset. Fig. 7 shows that for zero chip input, if a pair of adjacent comparators have different offsets, then each comparator outputs pulsewidths slightly different from the ideal $(1/2)T_{pwm}$. This causes the XOR gate to output nonzero pulsewidths proportional to the absolute difference between the two offsets, $|V_{os1} - V_{os2}|$, despite zero chip input. Therefore, the stopband rejection of the interpolated features is limited by comparator offset. A guideline is to limit

the comparator offset to a small fraction of the triangle-wave amplitude V_{tri}^{pp}.

That the stopband rejection of the interpolated features is limited by comparator offset is because at any frequency in the stopband of an interpolated feature, the corresponding bandpass filter is outputting nearly zero voltage, as if the chip input were zero. And, so, an interpolated feature's stopband value is instead given, up to a proportionality constant, by the interpolated feature's corresponding green XOR gate's zero-chip-input–output pulsewidth, which itself is a function of the comparator offset.

1) Calibration: Instead of increasing the area of the comparator by $M\times$ to reduce its offset by $\sqrt{M}\times$—because doing so would incur an $M\times$ power penalty to maintain the same delay—the $M\times$ larger area can be used instead to realize M copies of the comparator, of which the lowest-offset one can be selected with the other $M-1$ powered off. This technique offers two simultaneous advantages. The first is that the offset of the selected comparator is reduced by more than $\sqrt{M}\times$; for example, our MATLAB Monte Carlo simulations show that for $M = 8$, the reduction is $1.84\times$ larger[2] than the $\sqrt{8}\times$ reduction that the alternative $M\times$-larger-area comparator would have. The second, and more significant advantage, is that the selected comparator consumes $M\times$ less power than the alternative $M\times$-larger-area comparator would for the same delay. The concept of "selection" versus "averaging" of comparators was demonstrated as many as 20 years ago, that time in the application of flash ADCs [20].

Although it is not the standard deviation of a single comparator's offset, but rather the expected value of the absolute difference between the two offsets in a pair of comparators, $\mathbb{E}[|V_{os2} - V_{os1}|]$, that is sought to be reduced, selection can still be applied. In particular, in this chip, each of the $N = 16$ comparators is composed of $M = 8$ sub-comparators, and for each of the $N - 1 = 15$ pairs of adjacent comparators, that pair of sub-comparators that minimizes the interpolated feature output for zero chip input, is selected. This selection of sub-comparators is performed once for each chip, constituting a calibration that compensates for comparator offset to improve the stopband rejection of the interpolated features.

The calibration algorithm, which is executed OFF-chip by a personal computer (PC), entails: 1) feeding the chip zero input; 2) for a particular pair of adjacent comparators, programming the chip's scan chain to sweep across the 8×8 possible sub-comparators and, for each sweep point, measuring the corresponding interpolated feature chip output; 3) identifying the sub-comparator pair that minimizes this output; 4) repeating this sweep-and-identify for all pairs of adjacent comparators; and 5) re-programming the chip's scan chain such that all identified sub-comparators are selected.

We note that a viable alternative calibration technique for compensating for the comparator offset is to make the comparator's input device widths programmable.

[2]The 95% confidence interval for this factor is [1.82, 1.87], as calculated from a 10 000-trial MATLAB Monte Carlo simulation.
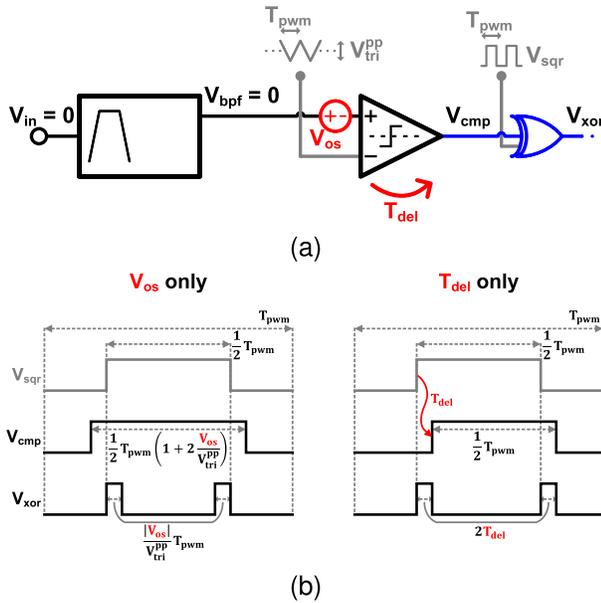
Fig. 8. Dominant non-idealities in time-mode analog rectification are comparator offset and delay (a), which cause the XOR gate to output nonzero pulsewidths (b) despite zero chip input. This limits the stopband rejection of the native features. The rejection can be improved by trimming the duty cycle and delay of the squarewave V_{sqr} to match those of the comparator output V_{cmp}. Note that the effect of offset and delay are exaggerated for illustration purposes.

B. Time-Mode Analog Rectification

Given the selected sub-comparator, its offset still enters into time-mode analog rectification, as does its delay, which together are the dominant non-idealities in time-mode analog rectification. Fig. 8 shows that comparator offset by itself causes the comparator output $V_{\text{cmp}}(t)$ to have pulsewidths slightly different from $(1/2)T_{\text{pwm}}$, whereas the squarewave $V_{\text{sqr}}(t)$ it's XOR'd against has $(1/2)T_{\text{pwm}}$-wide pulsewidths. This causes the XOR gate to output pulsewidths proportional to the absolute offset, $|V_{\text{os}}|$, despite zero chip input. Fig. 8 further shows that comparator delay by itself causes the comparator output $V_{\text{cmp}}(t)$ to be delayed with respect to the squarewave $V_{\text{sqr}}(t)$ it's XOR'd against. This causes the XOR gate to output pulsewidths proportional to T_{del}, despite zero input. Therefore, the stopband rejection of the native features is limited by both comparator offset and delay. In addition to the guideline to limit comparator offset to a small fraction of the triangle-wave amplitude $V_{\text{tri}}^{\text{pp}}$, a guideline for comparator delay is to limit it to a small fraction of the PWM period T_{pwm}.

That the stopband rejection of the native features is limited by comparator offset and delay is because at any frequency in the stopband of a native feature, the corresponding bandpass filter is outputting nearly zero voltage, as if the chip input were zero. And, so, a native feature's stopband value is instead given, up to a proportionality constant, by the native feature's corresponding blue XOR gate's zero-chip-input–output pulsewidth, which itself is a function of comparator offset and delay.

1) Calibration: Fig. 8(b) implies that trimming the duty cycle and delay of the squarewave V_{sqr} to match those of the

comparator output V_{cmp} would reduce the XOR gate output pulsewidth. This trimming is performed once for each chip, constituting a calibration that compensates for comparator offset and delay to improve the stopband rejection of the native features.

The calibration algorithm, which is executed off-chip by a PC, entails: 1) feeding the chip zero input; 2) programming the chip's scan chain such that the sub-comparators identified from the previous calibration procedure detailed in Section IV-A1 are selected; 3) for a particular native feature channel, sweeping the duty cycle and delay of the OFF-chip-generated-square-wave feeding the blue XOR gate in the channel and, for each sweep point, measuring the native feature chip output; 4) identifying the duty cycle and delay that minimize this output; 5) repeating this sweep-and-identify for all native feature channels; and 6) programming the OFF-chip digital pattern generator to generate square-waves with the identified duty cycles and delays.

V. CIRCUIT DESIGN

Fig. 2 shows that the chip architecture has the following circuit blocks in the signal path: filterbank, continuous-time comparators, XOR gates, IAFs, and counters. With the XOR gates and counters implemented with standard cells of the high-V_T flavor and near-minimum size, so as to decrease static leakage power and dynamic power, and the design of the IAFs similar to that in [13] except that a charge pump is used to convert the duty cycle of its input PWM waveform into a current, the designs of the filterbank and continuous-time comparators will be discussed.

A. Filterbank

The filterbank consists of 16 bandpass filters with center frequencies, f_c's, exponentially distributed from 100 Hz to 4 kHz by biasing the filters with currents, I_b's, that are exponentially distributed from 25 pA to 1 nA. This leverages the fact that at the nA current levels the filters are biased at, they are in deep weak inversion where (g_m/I_d) is constant at $\approx 30 \text{ V}^{-1}$, implying that $g_m \propto I_b$, and further implying that $f_c \propto g_m$, provided $f_c \propto (g_m/C)$ and C is kept the same across the bank. The necessary auxiliary circuit block that generates the pA-to-nA-level exponentially distributed currents is based on [21].

1) Filter: Fig. 9 shows the schematic of the bandpass filter, which, because of its power- and area- efficiency, is based on the super-source-follower (SSF) biquad, originally reported at radio frequency (RF) in its lowpass form [22], and since then re-purposed for audio in its bandpass form [13]. Because the original SSF topology [13], [22] suffers from limited common-mode and power supply rejection due to lack of a tail current source, we add M_{tail}, as is done in [23]. Because the original SSF topology [13], [22] suffers from limited quality factor Q due to transistor output resistances, we introduce negative resistance via the cross-coupled pair (M_{3p}, M_{3m}) to boost it. The resulting expressions for the filter's Q and f_c

Fig. 9. Schematic of the bandpass filter, which has a differential input and differential output.

are given by

$$Q = \frac{\sqrt{\frac{g_{m1}}{C_1}\frac{g_{m2}-g_3}{C_2}}}{\frac{g_{m2}}{C_2} - \frac{g_{m3}}{C_1 C_2/(C_1+C_2)}}, \qquad f_c = \sqrt{\frac{g_{m1}}{C_1}\frac{g_{m2}-g_{m3}}{C_2}}.$$

By choosing $(C_1 C_2)^{1/2}$, $(g_{m1}g_{m2})^{1/2}$, $g_{m1} = g_{m2}$, $(C_2/C_1)^{1/2}$, $g_{m3} = (1/5)g_{m2}$, the output integrated noise is set to 0.13 mV$_{rms}$–0.32 mV$_{rms}$ across the bank, the f_c is set to 100 Hz–4 kHz across the bank, the output integrated noise is minimized, the quality factor is set to 3, and oscillations are avoided, all respectively.

Further, because in deep weak inversion, the (g_m/I_d) of a device is constant independent of its W and L, both are freed up as design variables. In particular, $W \cdot L$ is increased to make the f_c standard deviation small enough to ensure with high confidence that the sequence of f_c's across the bank is monotonic, and W/L is decreased to make the total D–S leakage current in the lowest-frequency filter 10% of its total bias current.

B. Continuous-Time Comparator

Fig. 10 shows the schematic of the continuous-time comparator. Its function is to compare the voltage-mode output of the filter against a voltage-mode triangle wave (generated ON-chip using the circuit from [24]) to produce a PWM waveform interpreted as a time-mode representation of the filter output. Because the comparison must be done continuously over time, the comparator must be continouous-time instead of discrete-time, and is therefore implemented as a multi-stage operational transconductance amplifier (OTA). Because the filter output and triangle wave are both differential, the comparator must have a differential difference input. Because the comparator must interface with a subsequent XOR gate, it must have a single-ended, rail-to-rail output. Therefore, the comparator's first stage is a differential difference input/differential output OTA, the second stage a differential input/single-ended output wide-swing OTA, and the third stage is a cascade of inverters.

As IV explains, each comparator is composed of eight sub-comparators, of which one is selected, while the other seven are powered OFF. The Monte-Carlo-simulated random offset of the sub-comparator is [12.1 mV, 16.0 mV] with 95% confidence (with the systematic offset an order of magnitude smaller at 0.8 mV), which is further reduced after selection. Because the delay of the sub-comparator limits the min/max pulsewidths at its output, its bias current is set such that its delay is a small fraction, less than 10%, of the 250 μs PWM period, at a value of 17 μs. For this bias current, the simulated input-referred noise is 0.3 mV$_{rms}$, turning out to be much smaller than the input-referred offset.

VI. EXPERIMENTAL RESULTS

The 65 nm low power (LP) CMOS chip's die photograph is shown in Fig. 11(a). The chip occupies 0.53 mm^2 of active area, with Fig. 11(b) showing that the dominant contributors are the filters then comparators. The chip consumes 80 nW from the 0.6 V analog and 0.4 V digital supplies combined, with Fig. 11(b) showing that the dominant contributors are the comparators then filters. The chip's nW-level power consumption is measured by inserting a MΩ-sized precision resistor between the OFF-chip supply voltage and chip's supply pin, then measuring a mV-level voltage across it to infer supply current and thus power consumption, with the block-level breakdown inferred from simulation.

The die is mounted in a 64-pin QFP inserted into a socket that is soldered onto a PCB. The chip's analog input is generated using an Agilent 33 220 A 20 MHz function generator, and converted from single-ended to differential on-board, before being fed to the chip. The chip's 31-feature × 8-bit digital output is brought out on eight pins owing to an on-chip serializer, and captured using an Arduino Due.

The chip's scan chain is controlled using another Arduino Due, and the chip's reference squarewave inputs are generated using a Digilent Digital Discovery pattern generator. All measurement equipment and devices are controlled by a PC. Each of the chip's analog sub-systems—filterbank, comparator grid, IAF array, triangle wave generator—receives its own uA-level bias current from OFF-chip that once ON-chip is divided down by approximately 1000x to the necessary nA level using the current divider circuit from [21].
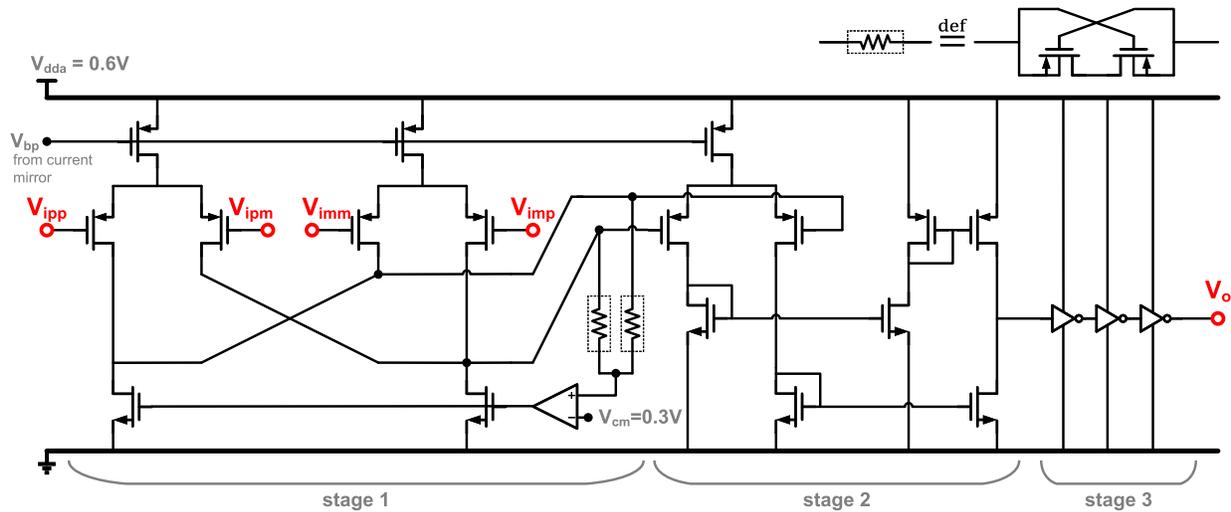
Fig. 10.   Schematic of the continuous-time comparator, which has a differential difference input and single-ended output.
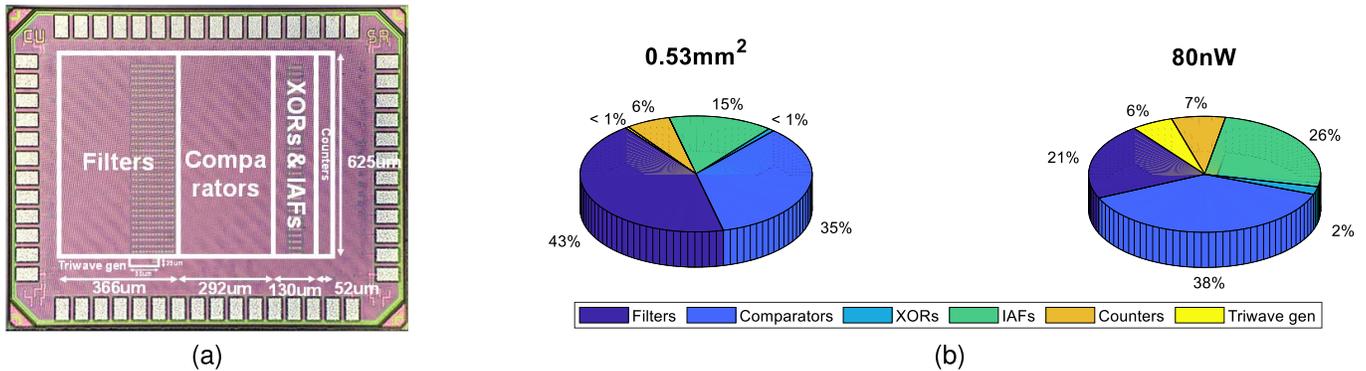

Fig. 11.   (a) Die photograph of the 65 nm LP chip. (b) Along with its area and power breakdowns.

A. Chip Characterization

The chip is characterized with a frequency response, as measured by feeding a differential analog sinewave into the chip, and capturing and averaging its 31 digital outputs. Fig. 12(a) shows the results for the 16 native features. Fig. 12(b) shows the results for 15 interpolated features, which, because they have bandpass shapes, confirms that time-mode analog filterbank interpolation indeed creates new bandpass features using the native features. Furthermore, Fig. 12(c) shows that the center frequencies of these new features fall in between those of the native features, confirming that these new features created by time-mode analog filterbank interpolation indeed have center frequencies that are interpolated. Fig. 12(c) additionally suggests that the native and interpolated center frequencies are well-controlled in that their sequence is monotonic. Even so, the interpolated center frequencies do slightly depart from their ideal positions at the geometric means of the adjacent pairs of native center frequencies. These departures are due to the amplitude variations in the native features.

Indeed, both the native and interpolated features exhibit variations in their passbands and stopbands. Their levels along with their causes are summarized in Table I. Such variations plague state-of-the-art analog audio feature extractors more

generally. For example, [8], [15], [17] exhibit 7.0, 7.1, and 8.2 dB passband variations, respectively, whereas in this work the passband variation is controlled to 5.4 dB. Despite their variations, [8], [15], [17] all report competitive KWS accuracies, as does this work to be shown in Section VI-B. This suggests that the backend classifiers used in all these works, this work included, are learning the variations in the analog audio feature extractors. The extent to which variation can be tolerated in the frontend analog feature extractor with the backend digital classifier learning the variation remains an open question, as does the dual question of the extent to which the backend can be simplified by reducing the variation in the frontend. Nevertheless, in this work, the variation could, in principle, be reduced further by calibrating the IAF gains through the IAF currents, as is done in [8].

B. KWS Demonstration

Having characterized the chip, it was then interfaced with a software backend classifier for a KWS demonstration. The Google Speech Commands (GSC) dataset was used [25]. Each example in the dataset is a pair consisting of a 1-s long, 8 kS/s, 16-bit audio clip of a keyword, along with the keyword label. We partitioned the dataset into training and test sets consisting
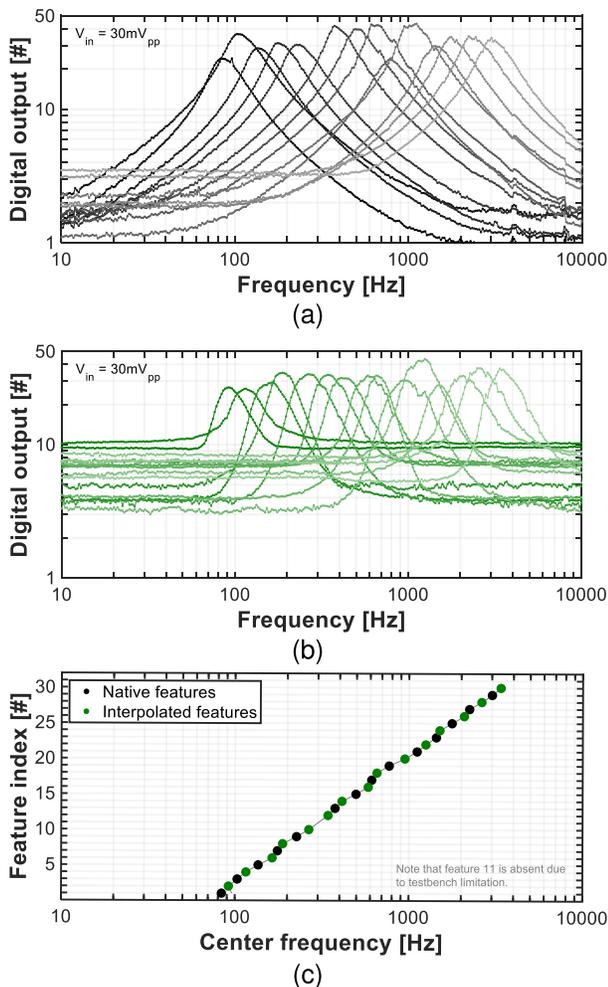
(a)

(b)

(c)

Fig. 12.   Chip characterization results. (a) Frequency responses of the native features. (b) Interpolated features. (c) Along with the progression of native and interpolated center frequencies.

TABLE I

SUMMARY OF LEVELS OF VARIATION IN NATIVE AND INTERPOLATED
FEATURES IN THEIR PASSBANDS AND STOPBANDS, ALONG WITH
THE CAUSES OF THE VARIATIONS

	Native		Interpolated	
Passband	5.4dB	Filter gain variation, IAF gain variation	4.5dB	Filter gain variation, IAF gain variation
Stopband	10dB	IAF gain variation	10dB	IAF gain variation, Comparator offset

of approximately 2500 and 350 examples, respectively, for a training-to-test ratio of 7:1. Each clip from the training set was generated as an analog signal by a function generator and fed into the chip, and the chip's resulting digital output spectrogram was captured, as shown in Fig. 13. The output spectrogram together with the keyword label constitutes one training example. This was repeated for all examples in the training set for two chips, and the same was done for the test set. The resulting two-chip training set was then used to train a standard small-footprint software neural network [26]. Finally, the two-chip test set was processed with the trained software
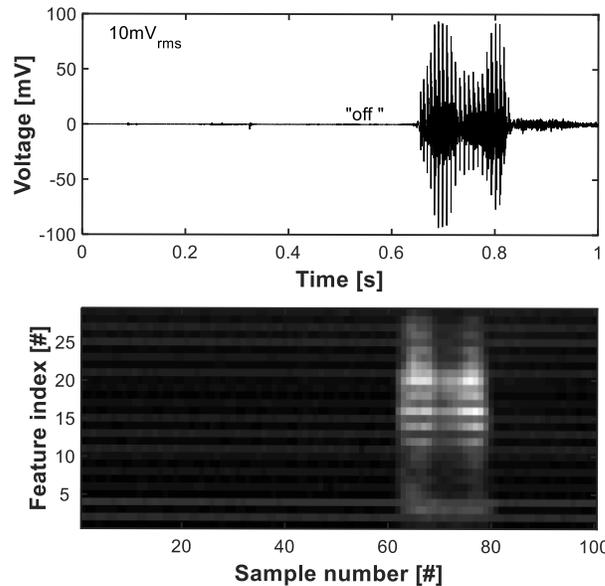

Fig. 13.   Example of analog audio chip input (top) and digital spectrogram chip output (bottom).
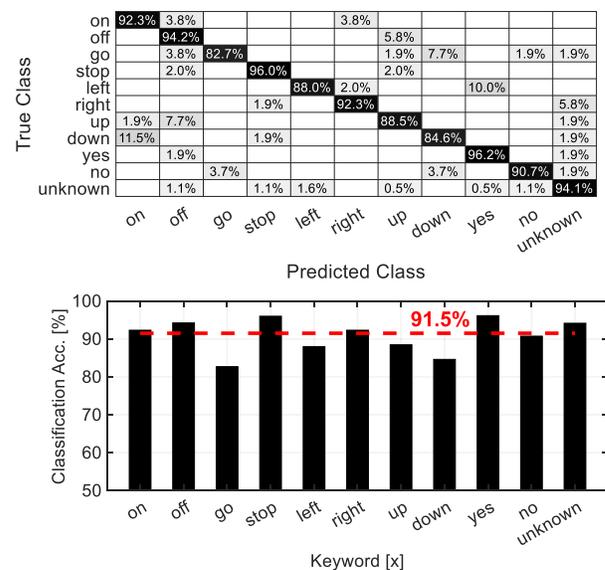

Fig. 14.   KWS demonstration results, summarized by a confusion matrix (top) and bar chart (bottom).

neural network to produce the confusion matrix and bar chart shown in Fig. 14. Even though the feature extractor exhibits significant variation (Table I), and the max/min in a typical spectrogram (Fig. 13) is modest at 35 dB, Fig. 14 shows that the classification accuracy is 91.5% across ten keywords—a competitive result that, as an indirect measure, proves the quality of the chip-extracted features.

To evaluate the marginal accuracy increase owing to the interpolated features, a software experiment was conducted. The interpolated features were removed from and native features retained in all spectrograms of the two-chip training and test sets. Then, the software neural network was retrained using the native-feature-only training set. Finally, the native-

Analog audio feature extractors

		10-keyword demo			< 10-keyword demo	
		This work	ISSCC'22 [17]	TCASI'21 [15]	ISSCC'21 [16]	JSSC'21 [8]
Signal processing technique	[x]	Time-Mode	Time-Mode	Switched-Capacitor	Nonlinear + Normalization	Nonlinear
Technology	[nm]	65	65	130	65	65
Input/output format	[x]	analog/digital	analog/digital[a]	analog/analog	analog/events	analog/events
Off-chip ADC required?	[x]	no	No	yes, to interface with classifier	no	no
Num of bandpass features	[#]	31	16	32	16	16
Power*	[nW]	80	11040[a]	800 (w/o off-chip ADC) 1500 (w/ off-chip ADC[b])	94	38
Area	[mm²]	0.53	1.6	0.79[c]	0.72	0.90
Power per feature	[nW]	2.6	690	25.0 (w/o off-chip ADC) 46.9 (w/ off-chip ADC)	5.9	2.4
Area per feature	[mm²]	0.017	0.1	0.025[c]	0.045	0.056
Frequency Range	[Hz - Hz]	100 - 4k	111 - 10.4k	30 - 8k	100 - 5k	100 - 5k
Bandwidth-normalized power per feature[†]	[nW]	9.5	1118	66.2 (w/o off-chip ADC) 124.1 (w/ off-chip ADC)	17.5	7.0
Feature rate	[Hz]	100	62.5	100	100	100.0

Keyword spotting demonstrations

Accuracy on GSC	[%]	91.5% on 10 words	86% on 10 words	82.4 - 87.4%[d] on 10 words	90.2% on 4 words	94.2% on 1 word
Keywords used	[x]	yes, no, up, down, left, right, on, off, stop, go	yes, no, up, down, left, right, on, off, stop, go	yes, no, up, down, left, right, on, off, stop, go	N/A	four
Classifier	[x]	off-chip software CNN	on-chip RNN	off-chip software RNN	off-chip hardware SNN	off-chip software CNN

*for fair comparison, power of mic pre-amp subtracted from [16], [8]

[†]calculated according to eqn used in [15]: $P_{\text{norm}} = P_{\text{total}} \cdot \frac{1-r}{1-r^N} \cdot \frac{4\text{kHz}}{f_{\max}}, \quad r = \left(\frac{f_{\min}}{f_{\max}}\right)^{\frac{1}{N-1}}$

[a]includes PFM decoder

[b]estimate of 700nW for off-chip ADC is from Fig. 15 in [15], [c]does not include area of necessary off-chip ADC

[d]includes the 5-10% degradation from excluding the off-chip logarithm, as stated in sec. IV-B of [15]

Fig. 15. Comparison table for state-of-the-art analog audio feature extractors with KWS demonstrations.

feature-only test set was processed with the native-feature-only-trained neural network, giving an accuracy of 90.8%. Together with the native-and-interpolated-feature 91.5% accuracy, this implies that the interpolated features increase the classification accuracy by 0.7% points, translating to a 7.6% improvement in classification error. This suggests that approaching 100% KWS accuracy becomes progressively challenging; for example, improving from 90% to 91% is likely more challenging than 89%–90%.

Nevertheless, the modest improvement calls into question the value of the interpolated features. Toward answering this question, a similar software experiment was conducted, but this time the native features were removed and the interpolated features retained. The resulting accuracy was 85.4%, which although not leading edge, is still above 80% and, in fact, is on par with the accuracy reported for the state-of-the-art switched-capacitor-based analog audio feature extractor chip plus software neural network reported in [15].

Taken together, we conclude that adding the interpolated features to the native features only modestly increases accuracy, although using the interpolated features alone does give adequate accuracy. Because this conclusion is specific to this particular chip's design point of $N = 16$ native features and $N - 1 = 15$ interpolated features, and keeping in mind that the interpolated features alone give adequate accuracy, we surmise that for a target accuracy of 90%, the hypothetical design point

of $N = 8$ native features and $N - 1 = 7$ interpolated features for a total $2N - 1 = 15$ features would be more power-optimal.

C. Comparison to State of the Art

Making fair comparisons among audio feature extractor chips is challenging for several reasons.

1) *There Is Little Consensus on the Type of Feature:* For example, [1] and [10] use Mel frequency cepstral coefficients (MFCCs), [13] uses Mel filterbank energies, [17] uses Mel filterbank energies with subsequent logarithm, and [16] uses Mel filterbank energies with a subsequent "per-channel energy normalization."

2) *Even for the Same Feature Type, There Is Little Consensus on its Parameters:* For example, for Mel filterbank energies, [27] uses ten filters ranging up to 2 kHz with quality factors as low as 0.7, while [15] uses 32 filters ranging up to 8 kHz with quality factors as high as 30.

3) *There Is Diversity in the Architecture of the Backend Classifier:* For example, [8] uses a convolutional neural network, [15] uses a recurrent neural network, and [16] uses a spiking neural network.

4) *There Is Diversity in the Recognition Task:* For example, [13] targets voice activity detection, [15] targets KWS, and [1] targets speaker identification (in addition to KWS).

With this context, toward a fair comparison, we narrow the focus to analog audio feature extractor chips with KWS demonstrations, as compared in Fig. 15. Compared to the most area-efficient previously published work [15], this work is $18\times$ for power-efficient in terms of power/feature, while compared to most power-efficient previously published work [8], this work is $3.3\times$ more area-efficient in terms of area/feature. And, compared to the only other work that uses time-mode ASP [17], this work is $265\times$ and $5.9\times$ more power- and area-efficient in terms of power/feature and area/feature.

The primary reason for the extreme $265\times$ power/feature gap between this work and [17]'s is that for the task of 10-word KWS, [17] targets unnecessarily high feature extractor dynamic range of 55 dB, whereas we target 35 dB. The secondary reason is that [17] targets unnecessarily high bandwidth of 10 kHz, whereas we target 4 kHz. Nevertheless, because the choices of dynamic range and bandwidth are task-dependent, and because given any analog audio feature extractor circuit architecture, its dynamic range and bandwidth can be scaled, a figure of merit (FoM) that normalizes power for dynamic range, bandwidth (as well as number of features) is warranted. Using the FoM created by [17], our work's FoM is 95.2 dB, which is 2.1 dB higher than [17]'s 93.1 dB FoM. However, the FoM created by [17] is not necessarily well-defined; for example, it rewards arbitrarily short frame shifts, although, for example, 1 ms frame shifts are known to be worse for speech recognition than 10 ms frame shifts. A well-defined FoM for analog audio feature extractors remains a challenging and open question.

VII. CONCLUSION

An analog audio feature extractor chip is reported that uses only 80 nW and 0.53 mm^2 to extract 31 features, advancing the state of the art in power/feature and area/feature by $18\times$ and $3.3\times$, as compared, respectively, to the most area- and power-efficient previously published analog audio feature extractor chips. All the while, competitive classification accuracy is maintained at $>90\%$ on ten keywords using a small-footprint software backend classifier. The chip's power- and area-efficiency is due to the proposed time-mode techniques, paving the way to broader adoption of the technology-scaling-favored time-mode ASP paradigm, in which analog information is encoded in the timing of digital edges, and processed by digital gates.

ACKNOWLEDGMENT

REFERENCES

[1] J. S. P. Giraldo, S. Lauwereins, K. Badami, and M. Verhelst, "Vocell: A 65-nm speech-triggered wake-up SoC for 10-μW keyword spotting and speaker verification," *IEEE J. Solid-State Circuits*, vol. 55, no. 4, pp. 868–878, Apr. 2020.

[2] N. C. Bui, J. J. Monbaron, and J. G. Michel, "An integrated voice recognition system," *IEEE J. Solid-State Circuits*, vol. SSC-18, no. 1, pp. 75–81, Feb. 1983.

[3] M. Verhelst and A. Bahai, "Where analog meets digital: Analog to information conversion and beyond," *IEEE Solid-State Circuits Mag.*, vol. 7, no. 3, pp. 67–80, Summer 2015.

[4] R. Sarpeshkar et al., "An ultra-low-power programmable analog bionic ear processor," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 4, pp. 711–727, Apr. 2005.

[5] R. Sarpeshkar, "Analog versus digital: Extrapolating from electronics to neurobiology," *Neural Comput.*, vol. 10, no. 7, pp. 1601–1638, Oct. 1998.

[6] E. A. Vittoz, "Future of analog in the VLSI environment," in *Proc. IEEE Int. Symp. Circuits Syst.*, May 1990, pp. 1372–1375.

[7] B. J. Hosticka, "Performance comparison of analog and digital circuits," *Proc. IEEE*, vol. 73, no. 1, pp. 25–29, Jan. 1985.

[8] M. Yang et al., "Nanowatt acoustic inference sensing exploiting non-linear analog feature extraction," *IEEE J. Solid-State Circuits*, vol. 56, no. 10, pp. 3123–3133, Oct. 2021.

[9] K. Badami, K. D. Murthy, P. Harpe, and M. Verhelst, "A 0.6 V 54 dB SNR analog frontend with 0.18% THD for low power sensory applications in 65 nm CMOS," in *Proc. IEEE Symp. VLSI Circuits*, Jun. 2018, pp. 241–242.

[10] L. Zhu, W. Shan, J. Xu, and Y. Lu, "AAD-KWS: A sub-μW keyword spotting chip with a zero-cost, acoustic activity detector from a 170 nW MFCC feature extractor in 28 nm CMOS," in *Proc. IEEE 47th Eur. Solid State Circuits Conf. (ESSCIRC)*, Sep. 2021, pp. 99–102.

[11] M. H. Najafi, S. Jamali-Zavareh, D. J. Lilja, M. D. Riedel, K. Bazargan, and R. Harjani, "Time-encoded values for highly efficient stochastic circuits," *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.*, vol. 25, no. 5, pp. 1644–1657, May 2017.

[12] S. Ray and P. R. Kinget, "A 31-feature, 80 nW, 0.53 mm^2 audio analog feature extractor based on time-mode analog filterbank interpolation and time-mode analog rectification," in *Proc. IEEE Symp. VLSI Technol. Circuits*, Jun. 2022, pp. 184–185.

[13] M. Yang, C.-H. Yeh, Y. Zhou, J. P. Cerqueira, A. A. Lazar, and M. Seok, "Design of an always-on deep neural network-based 1-μw voice activity detector aided with a customized software model for analog feature extraction," *IEEE J. Solid-State Circuits*, vol. 54, no. 6, pp. 1764–1777, Jun. 2019.

[14] Z. Song and D. V. Sarwate, "The frequency spectrum of pulse width modulated signals," *Signal Process.*, vol. 83, no. 10, pp. 2227–2258, 2003.

[15] D. A. Villamizar, D. G. Muratore, J. B. Wieser, and B. Murmann, "An 800 nW switched-capacitor feature extraction filterbank for sound classification," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 4, pp. 1578–1588, Apr. 2021.

[16] D. Wang, S. J. Kim, M. Yang, A. A. Lazar, and M. Seok, "A background-noise and process-variation-tolerant 109 nW acoustic feature extractor based on spike-domain divisive-energy normalization for an always-on keyword spotting device," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 64, Feb. 2021, pp. 160–162.

[17] K. Kim et al., "A 23-μW keyword spotting IC with ring-oscillator-based time-domain feature extraction," *IEEE J. Solid-State Circuits*, vol. 57, no. 11, pp. 3298–3311, Nov. 2022.

[18] H. Fuketa, "Ultralow power feature extractor using switched-capacitor-based bandpass filter, max operator, and neural network processor for keyword spotting," *IEEE Solid-State Circuits Lett.*, vol. 5, pp. 82–85, 2022.

[19] K. M. Badami, S. Lauwereins, W. Meert, and M. Verhelst, "A 90 nm CMOS, 6 μW power-proportional acoustic sensing frontend for voice activity detection," *IEEE J. Solid-State Circuits*, vol. 51, no. 1, pp. 291–302, Jan. 2016.

[20] M. P. Flynn, C. Donovan, and L. Sattler, "Digital calibration incorporating redundancy of flash ADCs," *IEEE Trans. Circuits Syst. II, Analog Digit. Signal Process.*, vol. 50, no. 5, pp. 205–213, May 2003.

[21] B. Linares-Barranco and T. Serrano-Gotarredona, "On the design and characterization of femtoampere current-mode circuits," *IEEE J. Solid-State Circuits*, vol. 38, no. 8, pp. 1353–1363, Aug. 2003.

[22] M. de Matteis, A. Pezzotta, S. D'Amico, and A. Baschirotto, "A 33 MHz 70 dB-SNR super-source-follower-based low-pass analog filter," *IEEE J. Solid-State Circuits*, vol. 50, no. 7, pp. 1516–1524, Jul. 2015.

[23] F. Fary, M. De Matteis, L. Rota, M. Arosio, and A. Baschirotto, "A 16 nm-FinFET 100 MHz 4th-order fully-differential super-source-follower analog filter," in *Proc. 26th IEEE Int. Conf. Electron., Circuits Syst. (ICECS)*, Nov. 2019, pp. 150–153.

[24] B. Vigraham, J. Kuppambatti, and P. R. Kinget, "Switched-mode operational amplifiers and their application to continuous-time filters in nanoscale CMOS," *IEEE J. Solid-State Circuits*, vol. 49, no. 12, pp. 2758–2772, Dec. 2014.

[25] P. Warden, "Speech commands: A dataset for limited-vocabulary speech recognition," 2018, *arXiv:1804.03209*.

[26] R. Tang and J. Lin, "Deep residual learning for small-footprint keyword spotting," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2018, pp. 5484–5488.

[27] E. Shi, X. Tang, and K. P. Pun, "A 270 nW switched-capacitor acoustic feature extractor for always-on voice activity detection," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 68, no. 3, pp. 1045–1054, Mar. 2021.
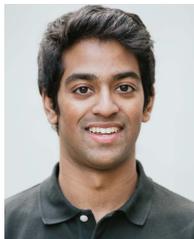
Subhajit Ray (Student Member, IEEE) received the B.S.E. degree in electrical engineering with a minor in physics from Princeton University, Princeton, NJ, USA, in 2015. He is currently pursuing the Ph.D. degree in electrical engineering with Columbia University, New York, NY, USA.

He has held a full-time position at the Air Force Research Laboratory, Rome, NY, and graduate internships at Analog Devices, Wilmington, MA, and SpaceX, Redmond, WA. His research interests are analog integrated circuits and systems, time-mode signal processing, analog and quantum computing, and machine learning applications.

Mr. Ray was a recipient of the Engineering Physics Senior Thesis Award in 2015, the CICC and ISSCC Student Travel Grant Awards in 2019 and 2022, respectively, and presented at the ISSCC 2022 Student Research Preview. He serves as a peer reviewer for the IEEE SOLID-STATE CIRCUITS LETTERS.

Peter R. Kinget (Fellow, IEEE) received the engineering degree (summa cum laude) in electrical and mechanical engineering and the Ph.D. degree (summa cum laude) in electrical engineering from Katholieke Universiteit Leuven, Leuven, Belgium, in 1990 and 1996, respectively.

From 1991 to 1995, he was a Research Assistant at the ESAT-MICAS Laboratory, Katholieke Universiteit Leuven. From 1996 to 1999, he was a member of Technical Staff with Bell Laboratories, Design Principles Department, Lucent Technologies, Murray Hill, NJ, USA. From 1999 to 2002, he held various technical and management positions in IC design and development at Broadcom, CeLight, and MultiLink. He is currently the Bernard J. Lechner Professor of Electrical Engineering with Columbia University, New York, NY, USA. He is widely published circuits and systems journal articles and conference papers. He has coauthored three books and holds 41 patents with several applications under review. His research interests include analog, radio frequency (RF) and power integrated circuits and the applications they enable in communications, sensing, and power management.

Dr. Kinget received a graduate fellowship from the Belgian National Fund for Scientific Research (NFWO) from 1991 to 1995. His research group has received funding from the National Science Foundation, the Semiconductor Research Corporation, the Department of Energy (ARPA-E), the Department of Defense (DARPA), and an IBM Faculty Award. It has further received in-kind and grant support from several of the major semiconductor companies. He is an inaugural recipient of the IEEE Solid-State Circuits Society Innovative Education Award in 2020. He was a co-recipient of several awards, including the Best Student Paper Award–1st Place from the 2008 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium, the First Prize from the 2009 Vodafone Americas Foundation Wireless Innovation Challenge, the Best Student Demo Award from the 2011 ACM Conference on Embedded Networked Sensor Systems (ACM SenSys), the 2011 IEEE Communications Society Award for Advances in Communications for an outstanding paper in any IEEE Communications Society publication in the past 15 years, the First Prize (U.S. $100K) from the 2012 Interdigital Wireless Innovation Challenge (I2C), the Best Student Paper Award–2nd Place from the 2015 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium, the Best Poster Award from the 2015 IEEE Custom Integrated Circuits Conference (CICC), and the Best Student Paper Award–3rd Place from the 2018 IEEE Radio Frequency Integrated Circuits (RFIC) Symposium. He served as a member for the Technical Program Committee of the IEEE Custom Integrated Circuits Conference from 2000 to 2005 and from 2016 to 2018, the Symposium on VLSI Circuits from 2003 to 2006, the European Solid-State Circuits Conference from 2005 to 2010, and the International Solid-State Circuits Conference from 2005 to 2012. He served as the Department Chair from 2017 to 2020. From 2010 to 2011, he was at the Université Catholique de Leuven, on sabbatical leave. He also serves as an expert for patent litigation and a technical consultant for industry. He was an IEEE Distinguished Lecturer of the Solid-State Circuits Society from 2009 to 2010 and from 2015 to 2017. He was an Elected Member of the IEEE Solid-State Circuits with Adcom from 2011 to 2013 and from 2014 to 2016, an Associate Editor of the IEEE JOURNAL OF SOLID-STATE CIRCUITS from 2003 to 2007 and the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS II: EXPRESS BRIEFS from 2008 to 2009.