This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

Modeling and Simulation of Circuit-Level Nonidealities for an Analog Computing Design Approach with Application to EEG Feature Extraction

Nikita Mirchandani, *Graduate Student Member, IEEE,* Yuqing Zhang, *Graduate Student Member, IEEE,* Safaa Abdelfattah, *Graduate Student Member, IEEE,* Marvin Onabajo, *Senior Member, IEEE,* Aatmesh Shrivastava, *Senior Member, IEEE*

***Abstract*—This paper presents a design approach for the modeling and simulation of ultra-low power (ULP) analog computing machine learning (ML) circuits for seizure detection using EEG signals in wearable health monitoring applications. In this paper, we describe a new analog system modeling and simulation technique to associate power consumption, noise, linearity, and other critical performance parameters of analog circuits with the classification accuracy of a given ML network, which allows to realize a power and performance optimized analog ML hardware implementation based on diverse application-specific needs. We carried out circuit simulations to obtain non-idealities, which are then mathematically modeled for an accurate mapping. We have modeled noise, non-linearity, resolution, and process variations such that the model can accurately obtain the classification accuracy of the analog computing based seizure detection system. Noise has been modeled as an input-referred white noise that can be directly added at the input. Device process and temperature variations were modeled as random fluctuations in circuit parameters such as gain and cut-off frequency. Non-linearity was mathematically modeled as a power series. The combined system level model was then simulated for classification accuracy assessments. The design approach helps to optimize power and area during the development of tailored analog circuits for ML networks with the ability to potentially trade power and performance goals while still ensuring the required classification accuracy. The simulation technique also enables to determine target specifications for each circuit block in the analog computing hardware. This is achieved by developing the ML hardware model, and investigating the effect of circuit nonidealities on classification accuracy. Simulation of an analog computing EEG seizure detection block shows a classification accuracy of 91%. The proposed modeling approach will significantly reduce design time and complexity of large analog computing systems. Two feature extraction approaches are also compared for an analog computing architecture.**

***Index Terms*—Analog computing, ultra-low power circuits, system modeling, feature extraction.**

I. Introduction

ELECTROENCEPHALOGRAPHY (EEG) is used for the diagnosis of neurological disorders such as epilepsy, sleep disorders, encephalopathies, and coma. In the case of epilepsy, treatment involves tracking and profiling of seizures to administer the correct medication. However, current treatment strategies, which includes interviewing patients or keeping them in the hospital for a long period, are either inaccurate or impractical [1]. For an accurate characterization of the onset of seizures, not only do we need to continuously monitor the EEG signal, but we also have to do it in an unobtrusive manner such that the day to day activities of a patient are not affected. A wearable device that can continuously monitor EEG suit this need. However, wearable battery-powered devices can monitor EEG only for a limited duration due to the relatively higher power consumption. In this paper, we propose an ultra-low power (ULP), analog EEG signal processing hardware and evaluate its capability through system-level modeling and simulations. The ULP operation of the hardware is based on robust sub-threshold analog computing circuits that are designed to process and extract a seizure event in the analog domain, incorporating ULP ML inference capability. The design approach introduced in this paper incorporates models based on non-idealities from circuit-level simulations in order to evaluate performance and power trade-offs during system-level simulations, which can be used to ease and expand the integration of analog computing components for ML.

A conventional EEG device converts the signal acquired by the analog front-end (AFE) [2] to a digital format using an analog-to-digital converter (ADC) followed by digital processing [3]. The data interface can be either wired or wireless [4]. While a wired data interface will station the patient close to a processing unit, a wireless EEG device can operate only for a limited duration due to the high power consumption. Wearable devices with wireless data communication [5] are in demand, but need lower power consumption. To that end, researchers have been developing on-chip capabilities for processing and feature extraction [6]–[9], which significantly reduces the amount of data transmission. In [7], [9], [10], seizure detection systems are introduced with power consumption in 100s of μWs. Another work reduces consumption down to $25\mu W$ [6]. In [11], the EEG feature extraction is performed directly on compressively-sensed signals to save power by processing fewer samples. Recently, an EEG monitoring device was reported that achieves a power consumption of 950 nW by employing an analog feature extraction technique [8]. In [12], a convolutional neural network (CNN) architecture is modeled

N. Mirchandani, Y. Zhang, S. Abdelfattah, M. Onabajo and A. Shrivastava are with the Department of Electrical and Computer Engineering, Northeastern University, Boston, MA, 02115 USA e-mail: (mirchandani.n@northeastern.edu; zhang.yuq@northeastern.edu; abdelfattah.s@northeastern.edu; monabajo@ece.neu.edu; aatmesh@ece.neu.edu).

This work is supported in part by National Science Foundation under grant CNS 1812588. (*Corresponding Author: Aatmesh Shrivastava*)

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

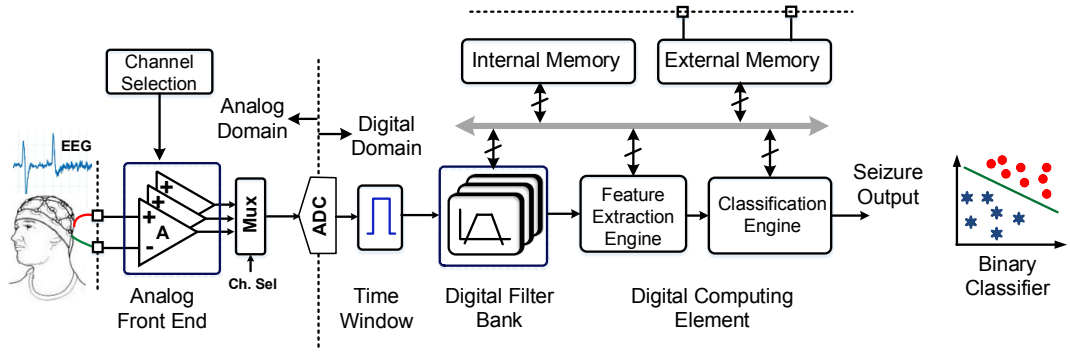

Fig. 1. EEG-based seizure detection system architecture using conventional digital computing.

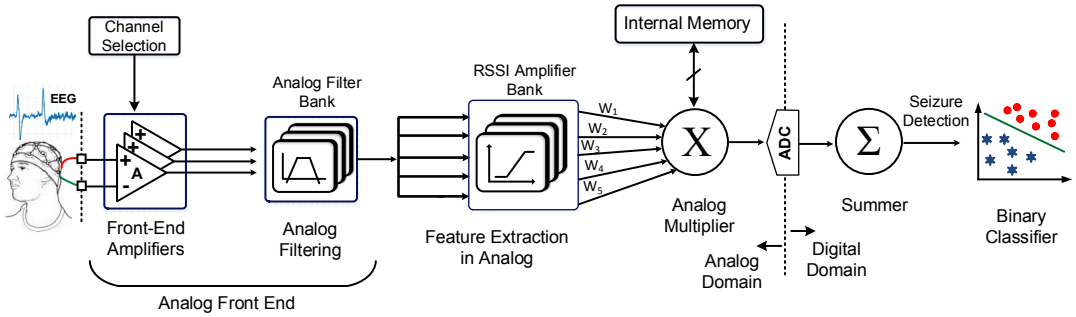

Fig. 2. Envisioned EEG-based seizure detection system architecture using analog computing.

to detect seizures with reduced computational cost. In [13], [14], deep learning algorithms are used to predict seizures.

In this paper, we demonstrate how ultra-low power analog computing circuits can be optimized with the added capability to consider their imperfections during system-level modeling and simulation aimed at ensuring high classification accuracy, where seizure detection is utilized as an example application. The complete signal processing chain of an EEG seizure detection system is implemented in the analog domain with only training weights stored in the digital domain. A new received signal strength indicator (RSSI) based feature extraction technique, better suited for analog computing, is also proposed. The analog computing approach not only saves power and area, but also has higher classification accuracy. Extensive circuit simulations were carried to accurately model various nonidealities in analog circuits such as noise, nonlinearity, and quantization error to assess their impacts on the classifier and to allow informed design decisions for underlying circuits. Conversely, in a bottom-up design approach, the modeling technique from this work can also be used to determine the system classification accuracy while exploring architectural alternatives based on the given performance parameters of each sub-circuit. Section II provides a background on EEG monitoring systems, while Section III introduces the proposed analog feature extraction circuits. Section IV describes the modeling of the analog system, and Section V explains the sources of nonidealities in each circuit block and how they are modeled. Section VI includes simulation results to exemplify the approach with feature extraction.

II. Background

Conventional seizure detection systems employ ML techniques to distinguish between seizures and normal EEG [15]–[19]. Normal EEG signals span the frequency range of $0-100$ Hz with spectrum being divided into five frequency bands: δ $0-3$ Hz, θ $3-7$ Hz, α $7-15$ Hz, β $15-32$ Hz, and $\gamma > 32$ Hz bands Specific frequencies become more prominent at the onset of seizures [20]. The power spectrum of the EEG signal is frequently used as a feature for classification. The frequency bands are differentiated using a filter bank of $4-7$ bandpass filters, and spectral energy of each sub-band is calculated with a moving time window. A feature map is constructed using this data and fed to a machine learning classifier.

Fig. 1 shows the architecture of a conventional EEG monitoring system. It includes an AFE, ADC, digital filters, feature extraction engine, and a binary classifier. At least 8 channels are used to acquire spatial information needed for high classification accuracy [21]. Often, ADCs with relatively high power consumption are also required to achieve high sampling rates to maintain signal fidelity. The 8 channel system presented in [7] utilizes an area of 25 mm^2, including the on-chip SVM classifier and a $64kB$ memory. In that, 8 channel AFE consumes $66\mu W$ of power.

EEG signals are at low frequencies and can be processed using ultra-low power, sub-threshold analog. Analog computing has reemerged as an alternate computing method to save power and area to realize a given function [22]–[30]. Recent works have also explored process variation compensation techniques to enhance reliability of analog neural networks [31].

In a digital system, circuit level results are abstracted out into boolean logic. Once abstracted out, circuit-level nonidealities are rarely investigated for their impact on system performance. This is not the case for analog computing, where circuit nonidealities can play a significant role in system performance. In the context of seizure detection using analog computing, a better understanding of performance impact of

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

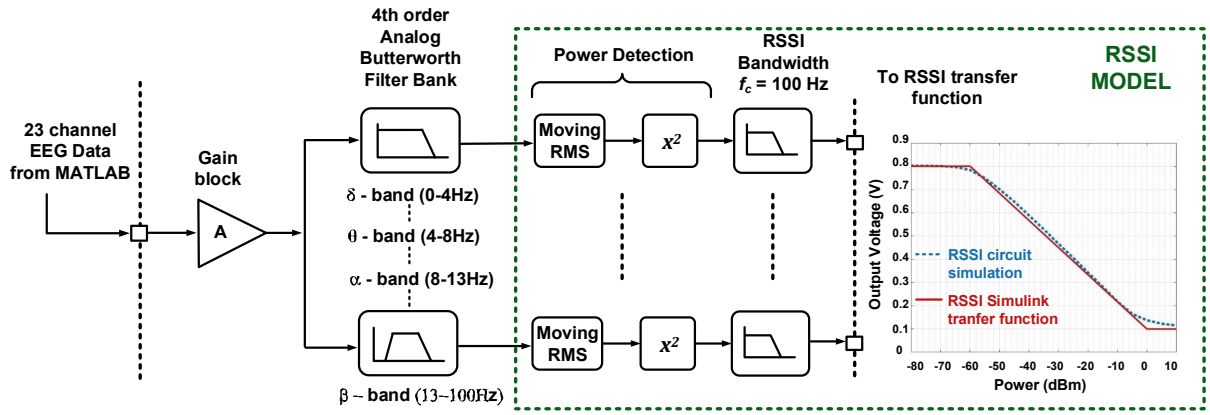

Fig. 3. Overview of the model components of the analog seizure detection architecture.

various blocks in the feature extraction unit is needed. The remainder of this paper will elaborate how an accurate analog model of the feature extraction engine can be utilized in the design process of fully analog computational units.

III. Analog Computing System Architecture

Fig. 2 displays an envisioned architecture of the analog processing system with analog feature extraction circuits. The first stage is composed of front-end amplifiers to provide gain for the incoming signal, which is followed by an analog filtering stage. The output of each filter is applied to a corresponding feature extraction circuit to continuously monitor for seizure. Various blocks with high-level functionality and specification in Fig. 2 are briefly discussed below.

1) Front-End Amplifiers: The amplifiers in the AFE boost the incoming EEG signal and should have a flatband gain of 20-30 dB, with ultra-low power consumption. They also have to be process-voltage-temperature (PVT) resilient.

2) Filter Bank: Filters for EEG have to be narrow band and require a sharp cut-off to reliably split the incoming signal into spectral bands. Filters suitable for biomedical applications with third-order harmonic distortion (HD_3) ranging from -60--40dB [32]–[34] are often used.

3) Feature Extraction Circuits: We explore two feature extraction techniques. The conventional root-mean-square (RMS) spectral energy of the frequency bands and a received signal strength indicator (RSSI) are modeled to compare their classification accuracy.

4) Analog-to-Digital Converter: The RSSI values are sent to an ultra-low power ADC. Since the ADC is moved to the end of the signal chain with the presented system architecture, the required sampling rate is very low at 4-samples/s for each channel, i.e., one sample per second for each of the four RSSI outputs. At such a low sampling rate, the power overhead of the ADC can be kept small.

5) Binary Classifier: The choice of classifier is application-dependent. Here, we have used a SVM classifier for binary classification. The SVM classifier consists of multiply and accumulate (MAC) units. The SVM classification computation can be implemented using off-chip or on-chip techniques [7], [35].

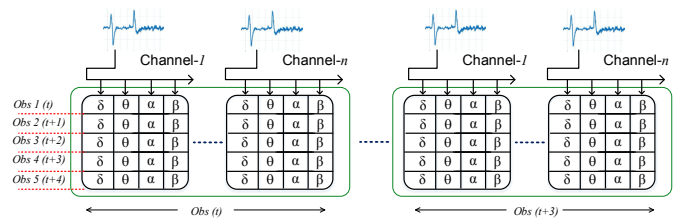

Fig. 4. Feature vector map creation using spectral power in a signal composed of n-channel EEG inputs.

IV. Analog Computing System Model

We constructed the system level architecture in Fig. 2 to obtain classification accuracy of seizures. Simulation times for the design of such complex analog systems can be prohibitively long, particularly when the impact of nonidealities such as noise, nonlinearity, and device mismatches have to be considered. We can reduce the simulation time by building accurate models. Furthermore, to facilitate the design of an analog computation engine, a study of the dependence of classification accuracy on circuit nonidealities is needed. The analog models also help define the required specifications to realize a given classification accuracy. First, an ideal model with the circuit blocks was created (Fig. 3) to obtain a baseline classification accuracy. Afterwards, nonidealities present in analog circuits were obtained using SPICE simulation and incorporated into the model to study their impact on ML classification accuracy. The system level architecture in Fig. 2 is modeled using behavioral level circuit functionality. For example, a linear gain element is used for the front-end amplifiers and filters that are modeled using real-time filter implementations. Similarly, the RSSI model is developed by closely mapping the input-output characteristics of an RSSI. Another important component of the model is feature definition and feature model creation using the EEG data, which is discussed below.

A. Feature Map Creation

Modeling of ML systems helps to develop better algorithms and hardware [36], [37]. The architecture introduced in Section III provides the training data for seizure classification. Next, a feature map must be prepared, which can be used as an input to a classifier. The accuracy of the classifier depends on the selected features. The EEG map must contain spectral, temporal, and spatial information associated with the seizures.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

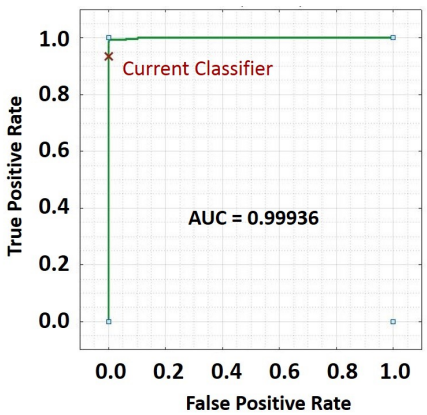

Fig. 5. ROC curve with AUC = 0.9993, showing that the two classes are separable.

The spectral components are obtained through the 4-stage filter bank. Spatial information is obtained by using multiple EEG channels. Seizures can be focal or general, and using more channels provides information about where a particular seizure originated in the brain. Motion and physiological artifacts, which corrupt EEG signals, can also be removed by increasing the number of channels. Usually, between 4 and 256 electrodes are used in present day EEG monitoring systems. Some applications also use channel selection algorithms to bring down the amount of data needed to be processed [38]. In this project, we use all 23 channels from the Boston Children's Hospital epilepsy database [21], [39]. The feature vector, X contains 92 features from the filter bank and 23 channels. The remaining temporal information is added by appending four feature vectors X calculated at one second intervals, creating feature vector X_T given by

$$X_T = \begin{bmatrix} X_t & X_{t+1} & X_{t+2} & X_{t+3} \end{bmatrix}$$

This results in a total of 368 features. The observations are recorded at an interval of one second. The feature vector is created for each training sample. The y-axis of the feature map represents the different training samples, which are each taken at one second intervals. The feature map is patient-specific and is fed to the linear SVM classifier. Fig. 4 shows a representation of the created feature map.

B. Model of the Analog System

Feature extraction block must be designed to achieve high seizure detection accuracy. We have modeled the blocks in Simulink to abstract out key design characteristics. First, an ideal model of the blocks was built based on the circuit level simulation of the analog block. The ideal model does not include nonidealities such as noise or nonlinearity. The nonideal characteristics were then introduced one at a time to study their effects on classifier performance. Fig. 3 visualizes the model blocks of the complete analog feature extraction engine. The EEG data from the CHB MIT database [39] is used, which is obtained at a sampling rate of 256 Hz. The simulation time step is also set to match the sampling rate. The gains of the filter and amplifier stages are modeled together. The gain block provides a flatband gain of 30dB, while the filter models include bandwidth restrictions. The raw EEG data is fed to the gain stage prior to four fourth-order analog butterworth filters to split the data into δ, θ, α, and β bands. The high cut-off frequencies of the filters are set to 4Hz, 8Hz, 12Hz, and 100Hz to cover the entire spectrum of the EEG data. The bandpass filters have been modeled as 4th-order filters with -40dB roll-off. The next block in the signal chain is the feature extraction engine. Both RMS and RSSI are modeled as part of the feature extraction engine. Since the incoming signal is continuous, the power is calculated in a 2 second time window. The average power of the signal is:

$$P_{avg} = \frac{1}{N}\sum_{n=0}^{N-1} |x_n|^2 = x_{rms}^2 \tag{1}$$

which is the square of the RMS level of the signal. Hence, the power of the signal can be modeled as a moving RMS block followed by a squaring block. The RSSI is modeled in two parts: first, a power detector circuit is modeled, and then the RSSI is modeled as a transfer function, where the conventional RSSI circuit with negative slope is modeled. If an RSSI has a positive slope, then the model can be adjusted through a sign change. Since the computation is done entirely in the analog domain, the RSSI signal is not digitized and the analog voltage level represents RSSI. The moving RMS block uses a sliding window with a length of 2 seconds. The squaring block outputs the power of the signal in a 2 second window. The RSSI is also modeled to have a low-pass frequency response and a 3-dB cut-off frequency of 100 Hz. The RSSI output waveform modeled as a linear function between -60dB and 0dB. Beyond this range, the graph is modeled as a constant function with an output voltage of 0 or $800mV$. The proposed model of RSSI was verified with simulation result of an RSSI circuit. Fig. 3 also shows the close agreement of an RSSI circuit simulation with the model. Once the data passes through the model of the feature extraction block, a feature map is built as explained in Section IV-A, and is fed to an SVM classifier.

C. Training and Classification

Training is typically patient-specific, and the data for EEG without seizures was collected for 148 hours. The data has been validated using 5-fold cross-validation to avoid over-fitting [40], [41]. For a single patient, 31 hours of normal EEG data and 402 seconds of seizure data have been used for training.

The classification is performed using a linear SVM (LSVM) classifier in the Statistics and Machine Learning Toolbox. The SVM classifier determines a decision boundary to separate the seizures from the non-seizures. A higher number of features results in a hyperplane separating the feature vector space into two classes. The general principle of a SVM classifier is as follows

$$W^T \cdot X + \beta = \begin{pmatrix} > 0; Seizure \\ < 0; Normal \end{pmatrix}$$

where X is the power spectral density vector of each of a channel, W^T and β are patient-specific trained weight data.

Since we need to wait for 4 seconds to collect all temporal information, it is not possible to conduct the final SVM in the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

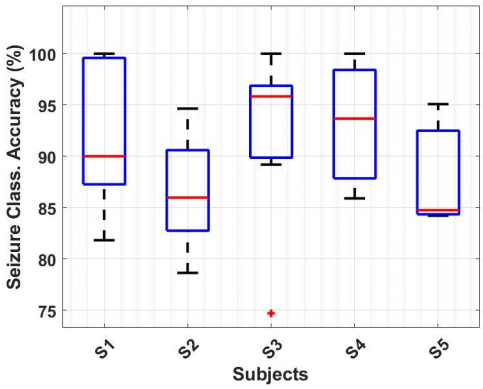

Fig. 6. Classification accuracy for seizure detection for 5 different patients.

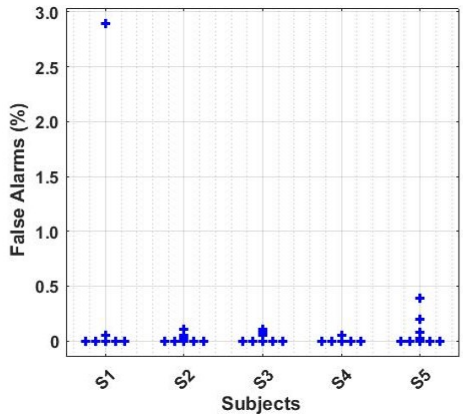

Fig. 7. Percentage of observations classified incorrectly as seizures for each patient. Each record was tested once, while the other records were used for training.

analog domain because that would require excessive analog memory. Storing the analog voltages for such a long time is currently not feasible due to the leakages in the switches. Instead, we perform $W_i \times X_i$ for each feature vector (i.e., the RSSI output) immediately when they become available, and the results are then converted into the digital domain for storage. At the end, when all temporal information is available, the saved values are summed to generate the final SVM output. Hence, the final summing for SVM is performed in digital, whereas the individual multiplications are performed in analog. Since analog-to-digital conversion is carried out at the end in one step, the sampling requirement for the ADC is relaxed significantly such that its power overhead can be low. The analog-to-digital conversion stage is required to include temporal information in 1s steps for 4s, and the outputs cannot be efficiently stored for such a long duration in the analog domain. The previous values have to be stored until all features are collected to create the feature vector. For applications where the temporal information is not required for feature map creation, the entire SVM operation can be performed in the analog domain.

Using this model, the classification accuracy was tested using the MIT CHB database. The feature map was formed using the ideal model explained above. The optimal number of spectral bands for RMS based feature extraction has been found to be 4 to 7 [7], [21]. The optimal number of spectral bands for RSSI based feature extraction is also 4 or higher as discussed in Section IV-D2. We have chosen and modeled 4 spectral bands for RSSI based feature extraction to keep the power consumption low. The receiver operating curve (ROC) is displayed in Fig. 5, where the area under the curve (AUC) is 0.99936.

The classifier is tested using patient-specific seizure data, and it outputs a class label for each observation of seizure data. The classification accuracy is calculated from the number of times the classifier correctly predicts each observation. The seizure classification accuracy, or the sensitivity is defined as

$$Classification\ Accuracy = \frac{TP}{TP + FN} \tag{2}$$

where TP (True Positives) is the number of seizure observations correctly predicted as seizures, and FN (False Negatives) is the number of seizure observations incorrectly predicted as no-seizure observations. The percentage of false alarms, or the specificity is defined as

$$False\ Alarms = \frac{TN}{TN + FP} \tag{3}$$

where TN is the number of True Negative observations, and FP is the number of False Positive observations. Since several records of seizures are available for each patient, each record is tested once while the other records are used for training. Fig. 6 shows the seizure classification accuracy for each seizure record of 5 different patients. The false alarm (FA) rate is also calculated in the same way, by cycling through records to test for each subject. Each observation falsely classified as a seizure is counted, and the total percentage of incorrectly labelled observations in a record is shown in Fig. 7.

D. Analog Feature Extraction Exploration

1) Comparison of Features: In a conventional EEG monitoring system, the spectral energy of the EEG signal sub-bands is used as a feature for seizure classification. The spectral energy in a sub-band is the RMS power of the EEG signal integrated over a time window. The difference between the RMS level in an EEG signal during seizure and normal brain activity forms the basis of classification. However, RMS power detection circuits are better suited for digital computation architectures. On the other hand, power feature extraction using RSSI can serve as a counterpart in analog. The RSSI is a log power detector, and is calculated using

$$RSSI_{OUT} = 10log(P_{in|dBm}) \tag{4}$$

The RSSI circuit consists of front-end amplifier stages to constitute a cascaded limiting amplifier structure for the realizing log function [42]. We model the RSSI as part of the feature extraction unit. The dynamic range of the RSSI presented in [29] was used for this application. The performance of RSSI and RMS power as features for seizure detection is compared in the following three ways. Firstly, the conventional digital architecture shown in Fig. 1 is used to extract RMS power. Four frequency bands are obtained using FIR filters dividing the incoming EEG signals into four equal sub-bands ranging from 0-30 Hz. High order filters are required to divide the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

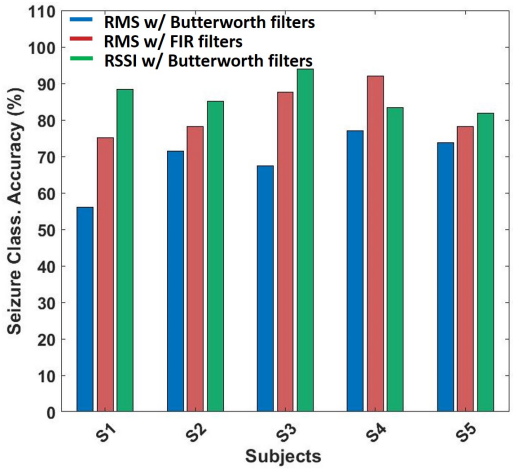

Fig. 8. Comparison of seizure classification accuracy with RSSI and RMS.

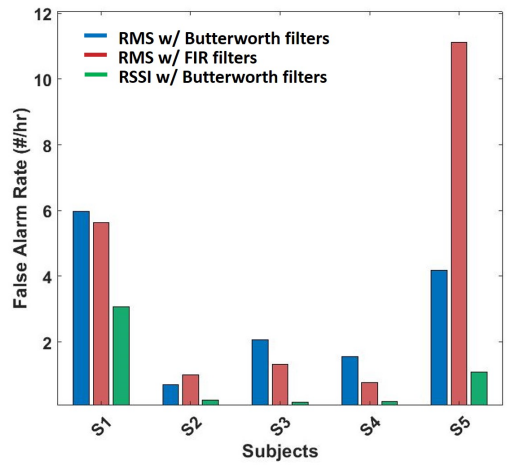

Fig. 9. Comparison of false alarm Rate with RSSI and RMS.

EEG signals into bands. We use 43rd-order FIR filters, as used in [7], in the filter bank.

In the second method, RMS is used as the feature extraction circuit using the analog system architecture shown in Fig. 2. The EEG signal is again divided into equal sub-bands from 0-30 Hz using fourth-order Butterworth filters.

In the third method, RSSI is used as the feature extraction circuit in the analog system architecture. The performance with both RSSI and RMS power is compared for seizure classification accuracy and FA rate.

Fig. 8 and Fig. 9 show the comparison between RMS and RSSI based feature extraction for five subjects. RSSI based feature extraction performs better across 5 subjects for both seizure classification accuracy and FA rate. Note that the difference in accuracy is small between FIR and RSSI, however, 40th-order FIR filters cannot be implemented in analog. The fourth-order RMS filters and RSSI are both feasible candidates for an analog computing system. Here, the difference between the RSSI and RMS methods was found to be statistically significant with $p = 0.0037$. The RSSI based feature extraction shows an average improvement of 17% in seizure classification accuracy across 5 patients, compared to the RMS based feature extraction. The number of FA per hour also improved by an average of 1.9 FA/hr for the RSSI based feature extraction compared to the RMS power based feature extraction. The better performance of RSSI is attributed to its linear response for a logarithmic input of amplified EEG signals to incorporate additional mathematical transformation. The RSSI circuit is also more compatible for an analog computing architecture, eliminating the need for corresponding higher-order filters.

2) Number of Spectral Bands: The amplified EEG signal is divided into frequency bands using band-pass filters. The three techniques of feature extraction are compared while sweeping the number of bands to find the optimal number of filters required for maximum accuracy and minimum false alarm rate.

The advantage of using an RSSI as the feature extraction circuit is further illustrated in Fig. 10. Here, the frequency range of interest (0 - 30 Hz) is divided into one to eight equal sub-bands for subject S1. The seizure classification accuracy increases with increasing number of filters for RMS power extracted using both FIR and analog filters. The RSSI shows consistently higher accuracy and smaller FA rate for all sub-bands. RSSI is also able to achieve higher accuracy with fewer number of filters. Hardware implementation of RSSI based feature extraction can hence save both power and area. In this paper, we use RSSI as the feature for seizure classification using four Butterworth filters. The bands are divided as δ (0-4Hz), θ (4 - 8Hz), α (8 - 13 Hz), and β (13 - 100 Hz) as these frequency bands give a slight performance boost compared to equally divided sub-bands.

V. Modeling of Block-Level Nonideal Characteristics and Simulation Results

The ideal model described in Section IV gives a baseline classification accuracy. In order to facilitate the design of the analog circuits within the feature extractor, nonidealities that can affect the classification accuracy should be taken into account. To achieve accurate classification results, circuit nonidealities were added to investigate their effects.

A. Noise

EEG signal processing is susceptible to noise due to the low amplitude of EEG signals. The electrodes, amplifiers, and filters all introduce an input-referred noise to the signals. If the noise is too high, the EEG signals can easily fall below the noise floor, resulting in a decreased classification accuracy. For this reason, biomedical signal acquisition systems generally have low input-referred noise [43]–[45] below 1 μV_{rms}.

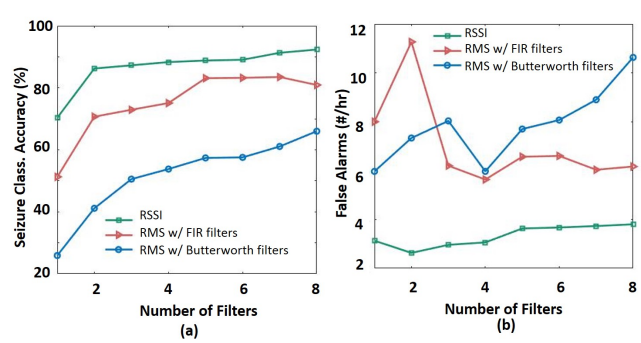

Fig. 10. Comparison of RSSI and RMS as features over a varying number of spectral bands.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

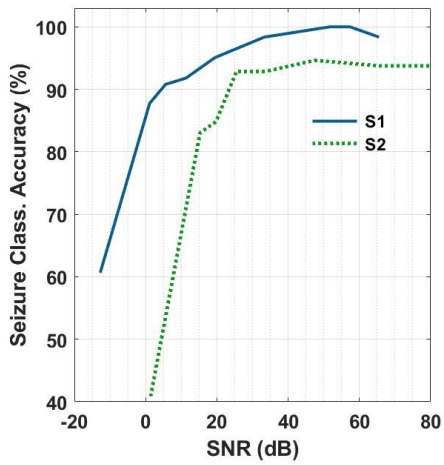

Fig. 11. Classifications accuracy vs. SNR for subjects S1 and S2, where the SNR is calculated using the seizure RMS value.

In this work, the ultra-low power, precise, sub-threshold front-end amplifier circuit discussed in [46] is modeled. Since this amplifier is switched capacitor based, it introduces high-frequency switching noise that is filtered out by the filter bank. The filter bank and RSSI also contribute to the input-referred noise, which are all lumped together as total input-referred noise.

To introduce noise in the model, normally distributed random numbers are generated with a specified standard deviation that is equivalent to the RMS value of the modeled input-referred noise. Noise is modeled as white noise in our MATLAB model, and has no dependence on frequency. This noise is then added directly to the EEG signals to account for the combined input-referred noise of the amplifier, filter, and RSSI. Note that this noise level was obtained through circuit-level noise and transient simulations with activated noise present in foundry-supplied devices models to accurately represent the impact of noise in our model.

The noise contributed by the AFE and RSSI was obtained by circuit level simulations. Fig. 11 shows the effect of noise on seizure classification accuracy for two different subjects, where the signal level is taken from the RMS of the seizure signals. Fig. 11 shows that the accuracy decreases exponentially for SNR$<$ 30dB.

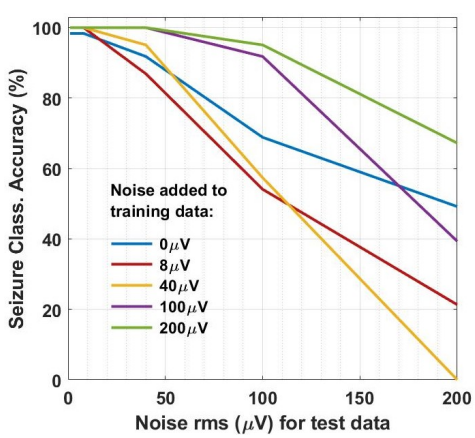

Fig. 12. Seizure classification accuracy with different levels of noise added during training for subject S1.

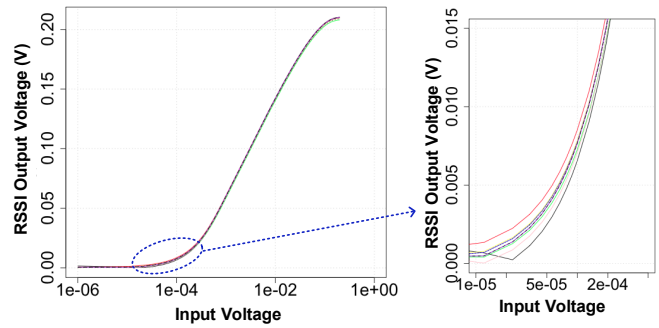

Fig. 13. RSSI output voltage simulated with transient noise. The maximum deviation of the RSSI output with noise is 4 mV.

Improving Robustness against Noise: In practical scenarios, the noise level varies between training and testing. To test the robustness of the system, different amounts of noise were added to the training and test data. Fig. 12 shows the classification accuracy versus added noise plots for subject S1. We varied the added noise from 0 to 200μV. The system is more robust when the training and test data is affected by the same level of noise. For example, when 200μV noise is added to the training of the EEG data, the system becomes robust against noise level lower than 200μV, and its classification accuracy does not degrade even when the noise amplitude is comparable to the seizure amplitude. Noise injection to inputs or weights during training improves regularization, and prevents overfitting [47]. However, there exists a trade-off between data fitting and regularization. Adding too much noise during training can make it harder for the model to fit the training data. The level of noise that the model can tolerate is patient-specific since the EEG signal amplitudes can vary greatly between patients. Fig. 12 shows that adding up to 200μV noise to the training data for subject S1 can have positive effect on the robustness of the model. To achieve high classification accuracy, the system design should ensure overall lower noise contributions from circuit components that can result in a higher power consumption. As the system noise starts to affect the classification accuracy, the model can be trained with added noise to improve classification accuracy.

B. RSSI Noise

The RSSI circuit also suffers from noise, which affects its sensitivity, especially at lower levels of power. When the signal power is low, it can fall below the noise level of the RSSI, resulting in a decreased dynamic range. Furthermore, the RSSI is impacted by device mismatches that create output offset. An offset correction technique is employed in each of the limiting amplifiers [29]. The effect of noise in the final RSSI output can be seen in Fig. 13. The offset correction scheme reduces the low frequency flicker noise. The overall noise is also reduced by increasing the output capacitor size. Corresponding to the power level of the input noise, the RSSI circuit will produce a DC voltage output following its transfer curve. Random variations in the noise power level in our integration time-window will generate the randomly distributed voltage level. This is shown in Fig. 13. This noise in the RSSI manifests

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

itself as an input-referred noise that can be referred back to the signal chain input. The average input-referred noise contribution (i.e., equivalent noise at the AFE input) from the RSSI block was found to be 0.1 μV, which is low, thanks to the gain in the signal path prior to the RSSI. The plots from Fig. 13 can be used to model the practical RSSI circuit. The RSSI noise results in variations in the output level of the RSSI at the low power levels, as shown by the simulation results in Fig. 13. The saturation voltage of the RSSI output varies by a maximum of 4 mV from the addition of noise, as verified by SPICE simulations. This variation in the output saturation level was modeled in the RSSI transfer curve. In our model, the output saturation level at low power levels was varied randomly using a uniform distribution ($\pm$ 4mV) for a conservative inclusion of the total RSSI output noise.

C. AFE Linearity

Differential analog filters and amplifiers predominately suffer from third-order distortion, since the second-order distortions are suppressed through the differential operation. In general, a memoryless nonlinear system produces the following output without offset:

$$y(t) = \alpha_1 x(t) + \alpha_2 x^2(t) + \alpha_3 x^3(t) \tag{5}$$

Here, α_1 is the small-signal gain of the system, and α_2 and α_3 are the second- and third-order distortion coefficients respectively. The AFE circuits are all fully differential, hence the second-order distortion is neglected in this analysis. The nonlinearity is modeled using a polynomial functional block. This introduces nonlinearity to the signal at the input of the filter bank. Higher order non-linearity coefficients are ignored, leaving the third-order harmonic distortion as the main nonlinear component. For small amplitude, the coefficient for the third order harmonic distortion is found using equation below, where A is amplitude of the signal

$$HD_3 = 20log\left(\frac{\alpha_3 A^2}{4\alpha_1}\right) \tag{6}$$

The HD_3 of the amplifier and filter bank blocks are combined

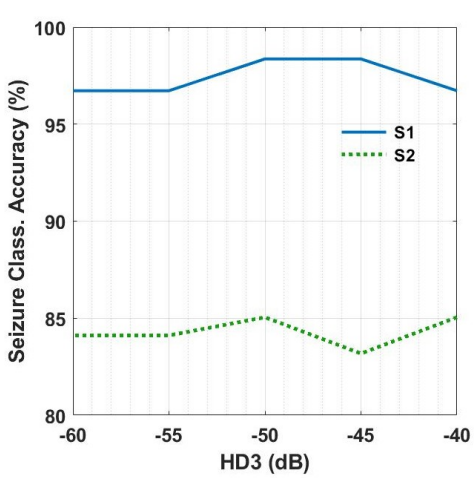

Fig. 14. Seizure classification accuracy vs. harmonic distortion for two subjects, S1 and S2.

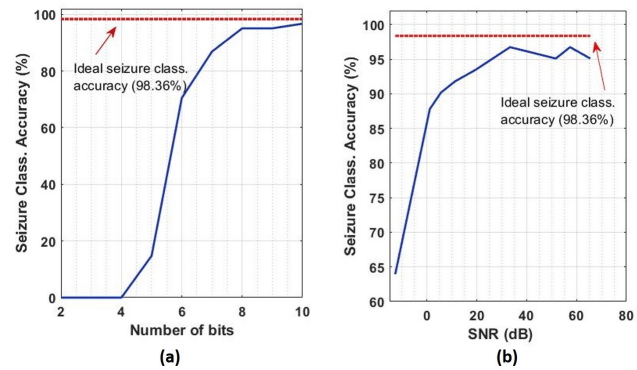

Fig. 15. (a) Seizure classification accuracy vs. resolution in number of bits for subject S1. (b) Seizure classification accuracy vs. SNR with weights quantized to 8 bits, showing the combined effects of noise and resolution on classification accuracy for subject S1.

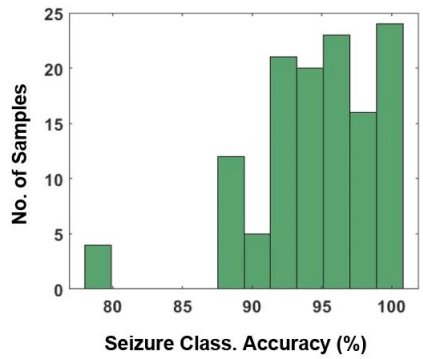

Fig. 16. Seizure classification accuracy with modeled PVT variation with σ = 10%

and modeled together. For EEG acquisition front-ends, the total harmonic distortion (THD) is usually kept low to around 1% [2], [48], [49]. Since signals are of low amplitude, the THD is a much looser constraint than noise. For a conservative estimate, the third-order coefficient is calculated using Eq. 6, with a 6mV input signal, which is a value that is higher than the expected amplified EEG signal. To ascertain how much nonlinearity the AFE can tolerate, the HD_3 was swept from -60 dBc to -40 dBc. Fig. 14 shows the seizure classification accuracy with respect to HD_3. The classification accuracy exhibits very low dependence on the linearity of the AFE. This is primarily because of two reasons: First, the power contribution due to nonlinearity is small, and second, the nonlinearity in the circuit does not change the relative power output produced by the RSSI circuit. The RSSI circuit will add both fundamental and HD_3 power together maintaining the relative ratio of power in the EEG spectral bands resulting in an insignificant impact on the classification accuracy.

D. Resolution

In standard SVM classifiers, floating point weights are used for training and inference. However, hardware implementation for real-time inference often requires fixed point resolution. Studies on low precision neural networks [50], [51] found that performance comparable to 32 bit floating point format could be achieved using quantized weights and activation. Quantized weights also reduce computing and storage requirements. The number of bits required for representation of weights is assessed in this section. Floating point weights obtained from

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

training are assumed to be quantized and stored on-chip before using them for inference. Fig. 15(a) shows the classification accuracy against number of bits used for quantization for subject S1. It reveals that the classification accuracy remains relatively high for a resolution of 7-bits or above. Here, we show the classification with the limited resolution of weights, but this essentially models the effective resolution of the DAC with multiplication and the RSSI circuit.

E. PVT Variations

PVT variations are factored into our model to account for deviations in the fabrication process. The gain and offset of the front-end amplifier are affected by the PVT variations. The most significant effect on the filter bank is cut-off frequency variation within each of the four filters. The random offset is corrected in the RSSI stage. However, the RSSI transfer function itself may vary due to PVT variations. This can modify both, the slope and the saturation voltages of the RSSI. The gain of the front-end amplifier, the cut-off frequencies of the filter bank, and the transfer function of the RSSI are all varied randomly around their means for a 1σ variation of 10%. The model is trained on ideal data without variations, and tested on data affected by PVT variations. The classification accuracy for 150 points is shown in Fig. 16. The random fluctuations are from a normal distribution, however, since the RSSI performs a non-linear transformation on the data, the seizure classification accuracy is not expected to follow a Gaussian distribution. In addition, the weights are assigned such that some features are weighted more than others, making the system non-linear. Note that using our sub-threshold biasing technique, a σ of 1% is seen with process variations, which does not show any significant variation in the classification accuracy.

F. System-level Simulation Results with Combined Effect of Nonidealities

The described analog based seizure detection system is robust against circuit nonidealities such as noise, nonlinearity, and quantization errors. The combined effect of these nonidealities is reported in this section. Fig. 15(b) shows the overall impact of noise and resolution on the seizure classification accuracy. The nonlinearity of the system is ignored since it has negligible effect on the sensitivity of the system. The ideal classification accuracy was obtained from the ideal model without quantization of weights. The noise added to both training and test datasets was varied, and all weights were quantized to 8 bits resolution. The classification accuracy shows minimal degradation for a SNR>30dB. For an ADC with $>$ 8-bit resolution, the system does not show significant degradation in the classification accuracy. Table I summarizes the simulated performance of the proposed approach. The latency of the detector was found to be 4 seconds, creating a delay between the onset and detection of seizures. The latency is a trade-off for the accuracy because 4 seconds of temporal information is used to create the feature map.

Table II shows the power and area breakdown of major circuit blocks of the system. The AFE and RSSI blocks were designed and simulated to estimate power consumption, which was found to be 268 nW per channel. Filter implementations for biomedical applications have been reported with power consumptions as low as a few nanowatts [32], [56], [57]. We used the multiplier cell from [58] with its area and power included in Table II. Since a multiplication output is required once every 1s, the multiplier is duty-cycled and its power is ≤ 0.02nW. 8-bit weights for the multipliers are stored in registers. Four such 8-bit registers are needed per channel before the signal is digitized. ADCs for biomedical applications consume low power due to their low sampling rates [55], [59]. For the analog computing system for EEG extraction, a sampling rate of only 4S/s is required. Hence, the power consumption of the ADC can be as low as $\sim$10nW based on [55]. The combined area of the system based on preliminary layout designs as well as estimation for the ADC [55] is $0.78mm^2$. We estimate the total area for one channel of the analog EEG feature extraction system to be $< 1.2mm^2$ while accounting for routing overhead and on-chip interfacing circuits. In comparison, the area for a single channel of digital implementations ranges from $3.125mm^2$ to $6.25mm^2$ [6], [7], [35], [52], [53]. The assessments of the analog feature extraction engine show that it can achieve high seizure classification accuracy with ultra-low power consumption. Furthermore, the total power consumption can be traded off with classification accuracy by decreasing the number of channels and frequency bands in the feature map.

Table I also summarizes the simulation results in comparison with other works. With the inclusion of the modeled circuit non-idealities, the described analog architecture maintains a similar system performance in terms of classification accuracy. While the system is tolerant to circuit-level distortions (non-linearities), the results have shown that the noise levels have to be kept in an acceptable range by design. Our analog computing ML model is also comparable to a recent CNN based model [12], where binarized inputs and weights achieved an AUC of 95.8%. The proposed analog model with ideal circuit model achieves a classification accuracy of 98.3%.

The presented modeling and simulation approach allows to determine circuit-level specifications to achieve application-specific accuracy goals while utilizing low-power circuits for energy-efficient processing. In addition, it can be used as a simulation tool to evaluate trade-offs among block-level specifications to optimize performance and power for analog computing systems on an application-specific basis.

VI. Design Case Studies

The model created for the feature extraction engine for seizure detection can be used to facilitate the design of the analog feature extraction hardware. The information gained from the models are used to ascertain specifications required for each of the AFE and feature extraction blocks to achieve the desired classification accuracy. In this section, the front-end amplifier and filter blocks for feature extraction are designed using the specifications acquired from the model.

A. Front-End Amplifier Design

Following the principles of sub-threshold biasing from [46], a fully differential amplifier is biased with a PTAT current

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

TABLE I
PERFORMANCE COMPARISON OF EEG FEATURE EXTRACTION METHODS.

	TBCAS'16 [6]	JSSC'13 [7]	TBCAS'16 [52]	JSSC'10 [35]	JSSC'15 [53]	TBCAS'11 [54]	This work*
Technology (nm)	180	180	180	180	180	130	**65**
Supply (V)	0.9	1.8	1.8	1	1.8	1.2	1
Classifier	Thresholding	LSVM	NLSVM	NLSVM	dual-LSVM	Phase Sync.	**LSVM**
Sensitivity (%)	98.5%	82.7%	95.1%	93%	95.7%	100	**91%**
False Rate/Specificity	4.4/hour	4.5%	0.94%	0.3 ± 0.7/hr	0.27/hr	1/hr	**0.025%**
# of Ch.	8	8	8	18	16	2	**23**
No. of frequency bands	4	7	7	7	7	5	**4**
Latency (secs)	9.1	< 2	2	6.7 ±3	1	-	**4**

*Simulated

TABLE II
PER CHANNEL POWER CONSUMPTION AND AREA OF CIRCUIT BLOCKS

	Power (nW)	Area ($\mu m \times \mu m$)
Front-end amplifiers	60	210 × 210
Filters	100	550 × 400
RSSI	96	4 ×430 × 230
Multiplier	0.02	135 × 135
ADC*	10	540 × 110
Digital/Registers/Clock Generation	1.5	250 × 270

Estimated based on [55], designed with 90 nm CMOS technology.

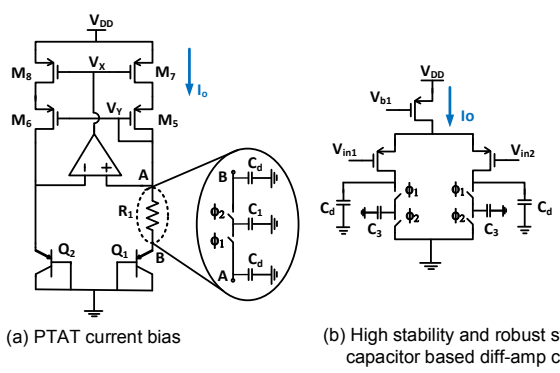

Fig. 17. (a) Constant g_m biasing, and (b) design of switched capacitor differential amplifier to realize robust operation [46].

source. The transconductance of the input pair transistors can be expressed as

$$g_m = \frac{ln(K)}{\eta R_1} \tag{7}$$

where R_1 is the resistance in the current source. The resistor is replaced with an equivalent switched capacitor resistor to ensure smaller design size. This biasing scheme can be used to ensure small transconductance variation, making analog circuits more robust against PVT variations.

The front-end amplifier is the first stage in the analog signal chain, which amplifies the incoming EEG signal. A fully differential amplifier with a gain of 22 dB was designed with the robust ultra-low power method described above, using the circuit in Fig. 17(b). This single-stage differential amplifier with resistive loads is biased with the described constant g_m biasing scheme. The load resistors are also replaced with equivalent switched capacitor resistors (SCRs). The gain with SCR load is given by

$$A_v = \frac{ln(K) f_1 C_1}{\eta f_3 C_3} \tag{8}$$

where f_3 is the switching frequency of the amplifier's SCR load, and f_1 is the switching frequency of the PTAT bias current generator. Equation 8 reveals the stable gain of the differential amplifier. The input amplitude of EEG AFEs is typically below 1mV as the MIT CHB data indicates [21], [39]. An amplifier circuit will exhibit better linearity performance for smaller amplitudes. For a conservative worst-case estimate, we simulated the amplifier for a maximum input voltage of 6mV with a 12Hz sinusoidal test signal to evaluate its nonlinearity. Fig. 18 displays the output spectrum generated by the circuit. The fractional third-order harmonic distortion (HD_3) component is -58dB below the fundamental component. The modeled nonlinearity is also shown in Fig. 18. The modeled nonlinearity closely matches the simulated HD_3 of the amplifier.

The model for the analog computing system indicates that the classification accuracy does not degrade for SNR $>$20dB, corresponding to a noise level of 80μV. The input-referred integrated noise as a function of the amplifier power

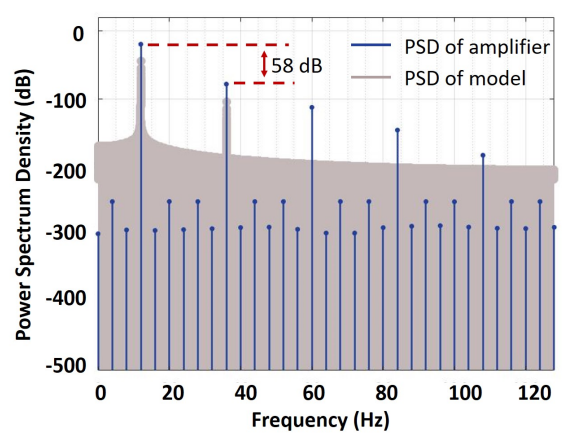

Fig. 18. HD_3 of the amplifier with an input voltage of 6mV, and frequency 12Hz.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

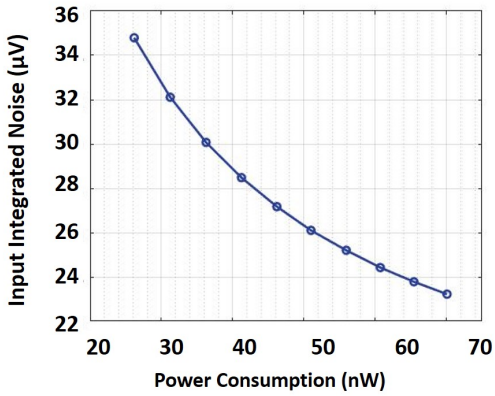

Fig. 19. Input-referred integrated noise of the amplifier as a function of power consumption.

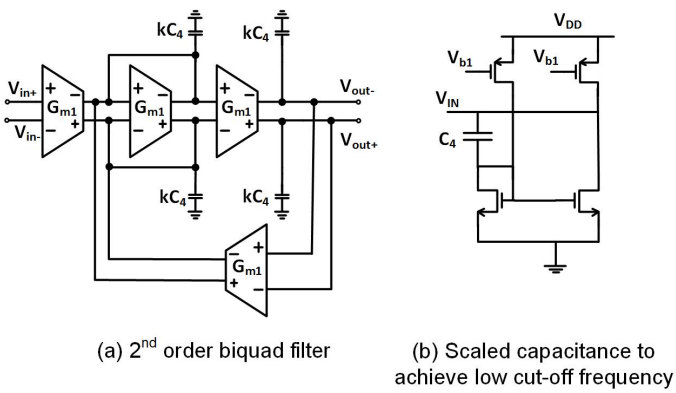

Fig. 20. (a) Second-order biquad filter. (b) Scaled capacitance to achieve low cut-off frequency with small area overhead.

consumption was simulated and is shown in Fig. 19; where an integration up to 10 kHz was performed, which is much higher than the AFE bandwidth. Input-referred noise was obtained from circuit level SPICE simulations by varying the bias current in the front-end amplifier while keeping the gain constant. The noise level is below the requirement, but note that the noise of the amplifier can be further decreased by increasing the power if needed.

The biasing circuit and the front-end amplifier both use switched capacitor circuits that require low-frequency clocks at 32kHz. The 32kHz clock can be generated with single digit nW power consumption [60]–[62]. A 32kHz crystal oscillator design based on [61] in 65 nm CMOS technology also produces single digit power consumption. Owing to the low operating frequency, the clock distribution will not incur significant power or routing area overhead

B. Filter Design

A second-order biquad filter was designed as shown in Fig. 20(a). The cut-off frequency of the filter is given by

$$\omega_o = \frac{g_m}{C_4} \tag{9}$$

Two biquads are cascaded together to obtain a fourth-order filter. The feature extraction engine requires low cut-off frequency (0-100 Hz) for the processing of EEG signals. This is achieved by lowering the transconductance of the OTA, and by

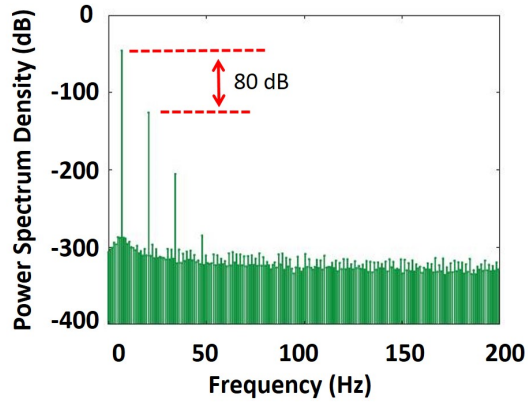

Fig. 21. HD_3 of the fourth order filter with an input voltage of 6mV, and frequency 6Hz. The input signal for linearity analysis is chosen to accurately represent the amplified EEG signal.

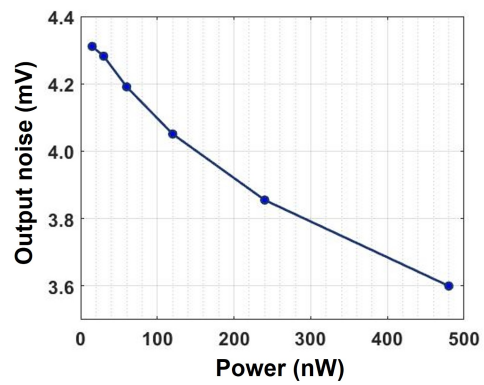

Fig. 22. Output noise of RSSI as a function of power consumption.

using large capacitors, which increase the area overhead. The transconductance is reduced by using a source degenerated OTA. A scaled capacitance, shown in Fig. 20(b), is used to scale up C_4 by $3\times$. The output noise of the filter was simulated while enabling transient noise using the foundry-supplied device models, and was found to be 180 μV integrated up to 1kHz. Hence, the combined input-referred noise of the front-end amplifiers and four filters is 83μV. As shown in Fig. 12, when the system is trained with noise, 83μV noise in the test set does not degrade the classification accuracy. The power consumption of the fourth-order biquad filter is 25 nW for a cut-off frequency of 8 Hz.

C. RSSI Design

A 6-stage RSSI similar to [29] was designed. Each limiting amplifier in the RSSI has a gain of 2.6. Transient output noise with zero input was simulated by varying the bias current. The resulting output noise as a function of power consumption is shown in Fig. 22. As expected, the RSSI output noise decreases with increasing power. However, the overall noise value when referred back to the input remains low. This helps during the informed scaling of the RSSI circuit power during the identification of design goals, while maintaining a high classification accuracy.

VII. Conclusion

In this paper, a robust analog computing architecture was introduced for EEG-based seizure detection. The main feature extraction components were designed and analyzed on the circuit level and modeled for system-level behavioral simulations. Furthermore, a comparison between two different features (RSSI and RMS) was performed to assess their impact on classification accuracy. A model of the feature extraction unit was created to study the effects of circuit-level nonidealities on seizure classification accuracy. The impacts of noise, nonlinearity and quantization of weights were examined. The AFE modeling and simulations showed that an SNR above 30dB ensures negligible effect on the classification accuracy with the described architecture. The results also revealed that the system robustness can be enhanced through training with an appropriate amount of added noise. The seizure detection rate under ideal conditions was 98.3%, whereas the detection rate with an SNR of 40dB and 8 bit quantization of weights was found to be 96%. The nonlinearity produced by the front-end amplifiers and filters had negigible impact on the accuracy of seizure detection. The average classification accuracy among all 5 subjects was found to be 91%, with a low average false positive rate of 0.025%. The power consumption of the AFE and feature extraction circuits was estimated to be 268 nW/channel.

References

[1] R. S. Fisher, D. E. Blum, B. DiVentura, J. Vannest, J. D. Hixson, R. Moss, S. T. Herman, B. E. Fureman, and J. A. French, "Seizure diaries for clinical research and practice: Limitations and future prospects," *Epilepsy and Behavior*, vol. 24, no. 3, 2012. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1525505012003368

[2] C. Chang, S. A. Zahrai, K. Wang, L. Xu, I. Farah, and M. Onabajo, "An Analog Front-End Chip With Self-Calibrated Input Impedance for Monitoring of Biosignals via Dry Electrode-Skin Interfaces," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 64, no. 10, 2017.

[3] T. Chen, H. Chiueh, S. Liang, S. Chang, C. Jeng, Y. Hsu, and T. Chien, "The implementation of a low-power biomedical signal processor for real-time epileptic seizure detection on absence animal models," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 1, no. 4, pp. 613–621, 2011.

[4] A. J. Casson, D. C. Yates, S. J. M. Smith, J. S. Duncan, and E. Rodriguez-Villegas, "Wearable Electroencephalography," *IEEE Engineering in Medicine and Biology Magazine*, vol. 29, no. 3, 2010.

[5] M. A. A. Ibrahim and M. Onabajo, "A Low-Power BFSK Transmitter Architecture for Biomedical Applications," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 67, no. 5, 2020.

[6] B. G. Do Valle, S. S. Cash, and C. G. Sodini, "Low-Power, 8-Channel EEG Recorder and Seizure Detector ASIC for a Subdermal Implantable System," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 6, pp. 1058–1067, 2016.

[7] J. Yoo, L. Yan, D. El-Damak, M. A. B. Altaf, A. H. Shoeb, and A. P. Chandrakasan, "An 8-Channel Scalable EEG Acquisition SoC with Patient-Specific Seizure Classification and Recording Processor," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 1, 2013.

[8] S. Iranmanesh and E. Rodriguez-Villegas, "A 950 nW Analog-Based Data Reduction Chip for Wearable EEG Systems in Epilepsy," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 9, 2017.

[9] M. Shoaib, N. K. Jha, and N. Verma, "A compressed-domain processor for seizure detection to simultaneously reduce computation and communication energy," in *Proceedings of the IEEE Custom Integrated Circuits Conference*, 2012.

[10] T. Roh, K. Song, H. Cho, D. Shin, and H. Yoo, "A Wearable Neuro-Feedback System with EEG-Based Mental Status Monitoring and Transcranial Electrical Stimulation," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 8, no. 6, 2014.

[11] M. Shoaib, N. K. Jha, and N. Verma, "Signal processing with direct computations on compressively sensed data," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 23, no. 1, pp. 30–43, 2015.

[12] N. D. Truong, A. D. Nguyen, L. Kuhlmann, M. R. Bonyadi, J. Yang, S. Ippolito, and O. Kavehei, "Integer convolutional neural network for seizure detection," *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 8, no. 4, pp. 849–857, 2018.

[13] H. Daoud and M. A. Bayoumi, "Efficient epileptic seizure prediction based on deep learning," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 5, pp. 804–813, 2019.

[14] Y. Xu, J. Yang, S. Zhao, H. Wu, and M. Sawan, "An end-to-end deep learning approach for epileptic seizure prediction," in *IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2020, pp. 266–270.

[15] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Principal Component Analysis-Enhanced Cosine Radial Basis Function Neural Network for Robust Epilepsy and Seizure Detection," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 2, 2008.

[16] T. Zhan, S. Z. Fatmi, S. Guraya, and H. Kassiri, "A Resource-Optimized VLSI Implementation of a Patient-Specific Seizure Detection Algorithm on a Custom-Made 2.2 cm^2 Wireless Device for Ambulatory Epilepsy Diagnostics," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 6, 2019.

[17] Siuly, Y. Li, and P. P. Wen, "Clustering technique-based least square support vector machine for EEG signal classification," *Computer Methods and Programs in Biomedicine*, vol. 104, no. 3, 2011. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0169260710002907

[18] A. K. Tiwari, R. B. Pachori, V. Kanhangad, and B. K. Panigrahi, "Automated Diagnosis of Epilepsy Using Key-Point-Based Local Binary Pattern of EEG Signals," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 4, 2017.

[19] H. Daoud and M. Bayoumi, "Deep Learning Approach for Epileptic Focus Localization," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 2, 2020.

[20] J. Gotman, J. R. Ives, and P. Gloor, "Frequency content of EEG and EMG at seizure onset: Possibility of removal of EMG artefact by digital filtering." *Electroencephalography and clinical neurophysiology*, vol. 52 6, 1981.

[21] A. Shoeb and J. Guttag, "Application of Machine Learning to Epileptic Seizure Detection," in *Proceedings of the International Conference on International Conference on Machine Learning*, ser. ICML'10. USA: Omnipress, 2010. [Online]. Available: http://dl.acm.org/citation.cfm?id=3104322.3104446

[22] B. Murmann, "Mixed-signal computing for deep neural network inference," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 29, no. 1, pp. 3–13, 2021.

[23] C. Popa, "Improved accuracy current-mode multiplier circuits with applications in analog signal processing," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 22, no. 2, pp. 443–447, 2014.

[24] C. Popa, "Improved accuracy pseudo-exponential function generator with applications in analog signal processing," in *The International Conference on "Computer as a Tool"*, vol. 2, 2005, pp. 1594–1597.

[25] J. Lu, S. Young, I. Arel, and J. Holleman, "A 1 TOPS/W Analog Deep Machine-Learning Engine With Floating-Gate Storage in 0.13 μm CMOS," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 1, 2015.

[26] J. Park, I. Hong, G. Kim, Y. Kim, K. Lee, S. Park, K. Bong, and H. Yoo, "A 646GOPS/W multi-classifier many-core processor with cortex-like architecture for super-resolution recognition," in *IEEE International Solid-State Circuits Conference Digest of Technical Papers*, 2013.

[27] Y. Huang, N. Guo, M. Seok, Y. Tsividis, K. Mandli, and S. Sethumadhavan, "Hybrid Analog-Digital Solution of Nonlinear Partial Differential Equations," in *IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 2017.

[28] M. R. Azghadi, B. Linares-Barranco, D. Abbott, and P. H. W. Leong, "A Hybrid CMOS-Memristor Neuromorphic Synapse," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 11, no. 2, 2017.

[29] Y. Zhang, N. Mirchandani, M. Onabajo, and A. Shrivastava, "RSSI Amplifier Design for a Feature Extraction Technique to Detect Seizures with Analog Computing," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, May 2020.

[30] D. Miyashita, S. Kousai, T. Suzuki, and J. Deguchi, "A Neuromorphic Chip Optimized for Deep Learning and CMOS Technology With Time-Domain Analog and Digital Mixed-Signal processing," *IEEE Journal of Solid-State Circuits*, vol. 52, no. 10, 2017.

[31] S. Moon, K. Shin, and D. Jeon, "Enhancing reliability of analog neural network processors," *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, vol. 27, no. 6, pp. 1455–1459, 2019.

[32] T. Zhang, P. Mak, M. Vai, P. Mak, M. Law, S. Pun, F. Wan, and R. P. Martins, "15-nW Biopotential LPFs in 0.35- μm CMOS Using Subthreshold-Source-Follower Biquads With and Without Gain Compensation," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 7, no. 5, 2013.

[33] S. Thanapithak and C. Sawigun, "A 1.5 V 5.2 nW 60 dB-DR Lowpass Filter With Self-Compensated Gain in 0.35 μm CMOS Suitable for Biomedical Applications," in *IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018.

[34] K. Wang, C. Chang, and M. Onabajo, "A fully-differential CMOS low-pass notch filter for biosignal measurement devices with high interference rejection," in *International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2014.

[35] N. Verma, A. Shoeb, J. Bohorquez, J. Dawson, J. Guttag, and A. P. Chandrakasan, "A micro-power eeg acquisition soc with integrated feature extraction processor for a chronic seizure detection system," *IEEE Journal of Solid-State Circuits*, vol. 45, no. 4, pp. 804–816, 2010.

[36] J. Acharya and A. Basu, "Deep Neural Network for Respiratory Sound Classification in Wearable Devices Enabled by Patient Specific Model Tuning," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 3, 2020.

[37] B. Zhu, M. Farivar, and M. Shoaran, "ResOT: Resource-Efficient Oblique Trees for Neural Signal Classification," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 14, no. 4, 2020.

[38] Huan Liu and Lei Yu, "Toward integrating feature selection algorithms for classification and clustering," *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 4, 2005.

[39] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R. Mark, and H. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals."

[40] T. Hastie, R. Tibshirani, and J. Friedman, *Elements of Statistical Learning: Data Mining, Inference, and Prediction*, ser. Springer series in statistics. New York, NY: Springer, 2009.

[41] J. Lever, M. Krzywinski, and N. Altman, "Model selection and overfitting," *Nature Methods*, vol. 13, September, 2016.

[42] J. Ramos, J. L. Ausín, J. F. Duque-Carrillo, and G. Torelli, "Design of limiting/logarithmic amplifier for wideband bioimpedance measuring devices," in *Biomedical Circuits and Systems Conference (BioCAS)*, Nov 2010.

[43] C. Wu and C. Ho, "An 8-channel chopper-stabilized analog front-end amplifier for EEG acquisition in 65-nm CMOS," in *IEEE Asian Solid-State Circuits Conference (A-SSCC)*, 2015.

[44] S. Lim, C. Seok, H. Kim, H. Song, and H. Ko, "A fully integrated electroencephalogram (EEG) analog front-end IC with capacitive input impedance boosting loop," in *Proceedings of the IEEE Custom Integrated Circuits Conference*, 2014, pp. 1–4.

[45] L. Dong, X. Gui, Y. Xie, and L. Geng, "An 8-Channel Biomedical Front-End (BFE) with 6.2-μW Per Channel for EEG/ECoG Acquisition," in *IEEE International Conference on Integrated Circuits, Technologies and Applications (ICTA)*, 2018.

[46] N. Mirchandani and A. Shrivastava, "High Stability Gain Structure and Filter Realization with less than 50 ppm$^\circ$/C Temperature Variation with Ultra-low Power Consumption using Switched-capacitor and Subthreshold Biasing," in *2018 IEEE International Symposium on Circuits and Systems (ISCAS)*, 2018.

[47] H. Noh, T. You, J. Mun, and B. Han, "Regularizing Deep Neural Networks by Noise: Its Interpretation and Optimization," in *Advances in Neural Information Processing Systems 30*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds. Curran Associates, Inc., 2017, pp. 5109–5118.

[48] T. Yang, J. Lu, M. S. Jahan, K. Griffin, J. Langford, and J. Holleman, "A configurable 5.9 μW analog front-end for biosignal acquisition," in *IEEE Custom Integrated Circuits Conference (CICC)*, 2015.

[49] Yu-Pin Hsu, Chun-Hao Chen, Hsiao-Chin Chen, and Shey-Shi Lu, "Low-noise wide-dynamic-range CMOS analog front-end IC for portable biomedical applications," in *Asia-Pacific Microwave Conference*, 2008.

[50] I. Hubara, M. Courbariaux, D. Soudry, R. El-Yaniv, and Y. Bengio, "Quantized Neural Networks: Training Neural Networks with Low Precision Weights and Activations," 2016. [Online]. Available: https://arxiv.org/abs/1609.07061

[51] Y. Guo, "A Survey on Methods and Theories of Quantized Neural Networks," 2018. [Online]. Available: https://arxiv.org/abs/1808.04752

[52] M. A. Bin Miyazaki and J. Yoo, "A 1.83 μJ/Classification, 8-Channel, Patient-Specific Epileptic Seizure Classification SoC Using a Non-Linear Support Vector Machine," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 10, no. 1, Feb 2016.

[53] M. A. Bin Altaf, C. Zhang, and J. Yoo, "A 16-Channel Patient-Specific Seizure Onset and Termination Detection SoC With Impedance-Adaptive Transcranial Electrical Stimulator," *IEEE Journal of Solid-State Circuits*, vol. 50, no. 11, 2015.

[54] K. Abdelhalim, V. Smolyakov, and R. Genov, "Phase-Synchronization Early Epileptic Seizure Detector VLSI Architecture," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 5, no. 5, 2011.

[55] S.-E. Hsieh and C.-C. Hsieh, "A 0.4V 13b 270kS/S SAR-ISDM ADC with an opamp-less time-domain integrator," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2018.

[56] S. Thanapitak and C. Sawigun, "A subthreshold buffer-based biquadratic cell and its application to biopotential filter design," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 65, no. 9, pp. 2774–2783, 2018.

[57] C. Sawigun and P. Pawarangkoon, "0.6-V, sub-nW, second-order low-pass filters using flipped voltage followers," in *IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, 2016.

[58] N. Mirchandani and A. Shrivastava, "CMOS based Ultra-low Power High-Precision Analog Vector Matrix Multiplication Circuit with $\pm 0.1\%$ error for vision application," in *IEEE International Midwest Symposium on Circuits and Systems (MWSCAS)*, 2019.

[59] S. Jeong, Y. Chen, T. Jang, J. Tsai, D. Blaauw, H.-S. Kim, and D. Sylvester, "21.6 a 12nW always-on acoustic sensing and object recognition microsystem using frequency-domain feature extraction and SVM classification," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2017.

[60] L. Xu, D. Blaauw, and D. Sylvester, "Ultra-Low Power 32kHz Crystal Oscillators: Fundamentals and Design Techniques," *IEEE Open Journal of the Solid-State Circuits Society*, vol. 1, pp. 79–93, 2021.

[61] A. Shrivastava, D. Akella Kamakshi, and B. H. Calhoun, "A 1.5 nW, 32.768 kHz XTAL Oscillator Operational From a 0.3 V Supply," *IEEE Journal of Solid-State Circuits*, vol. 51, no. 3, pp. 686–696, 2016.

[62] D. Yoon, D. Sylvester, and D. Blaauw, "A 5.58nW 32.768kHz DLL-assisted XO for real-time clocks in wireless sensing applications," in *IEEE International Solid-State Circuits Conference (ISSCC)*, 2012.

Nikita Mirchandani (S'17) received the B.E. (Hons.) degree in electrical and electronics engineering from Birla Institute of Technology and Science (BITS Pilani), India in 2015. She completed her M.S. in electrical and computer Engineering from Northeastern University, MA, USA in 2017. Currently, she is working towards the Ph.D. degree in electrical engineering at Northeastern University. Her research interests include the design of low power precision circuits and analog computing hardware for machine learning applications.

Yuqing Zhang (S'20) completed her bachelor's degree in electrical engineering at Shenzhen University, China in 2018. She is pursuing her master's degree at Northeastern University. Her research is related to low power IC design and analog computing circuits.

Safaa Abdelfattah (S'19) received the B.S. and M.Sc. degrees in electrical engineering from Cairo University, Giza, Egypt, in 2012 and 2016, respectively. She is currently pursuing the Ph.D. degree in the Department of Electrical and Computer Engineering at Northeastern University, Boston, MA, USA. She is working on research on mixed-signal design for low-power EEG signal acquisition. From 2012 to 2018, she was a Teaching and Research Assistant at Cairo University. Since 2018, she has been a graduate research assistant at Northeastern University.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TCAD.2022.3170248, IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems

Marvin Onabajo (S'01–M'10–SM'14) is an Associate Professor in the Electrical and Computer Engineering Department at Northeastern University. He received a B.S. degree (summa cum laude) in Electrical Engineering from The University of Texas at Arlington in 2003 as well as the M.S. and Ph.D. degrees in Electrical Engineering from Texas A&M University in 2007 and 2011, respectively.

From 2004 to 2005, he was Electrical Test/Product Engineer at Intel Corp. in Hillsboro, Oregon. He joined the Analog and Mixed-Signal Center at Texas A&M University in 2005, where he was engaged in research projects involving analog built-in testing, data converters, and on-chip temperature sensors for thermal monitoring. In the spring 2011 semester, he worked as a Design Engineering Intern in the Broadband RF/Tuner Development group at Broadcom Corp. in Irvine, California. Marvin Onabajo has been at Northeastern University since the Fall 2011 semester. His research areas are analog/RF integrated circuit design, on-chip built-in testing and calibration, mixed-signal integrated circuits for medical applications, data converters, and on-chip sensors for thermal monitoring. He currently serves as Associate Editor on the editorial boards of the IEEE Transactions on Circuits and Systems I (TCAS-I, 2016-2017, 2018-2019 and 2020-2021 terms) and of the IEEE Circuits and Systems Magazine (2016-2017, 2018-2019 and 2020-2021 terms). During the 2014-2015 term, he was on the editorial board of the IEEE Transactions on Circuits and Systems II (TCAS-II). He received a 2015 CAREER Award from the National Science Foundation, a 2017 Young Investigator Program Award from the Army Research Office (ARO), and the 2015 Martin Essigman Outstanding Teaching Award from the College of Engineering at Northeastern University.

Aatmesh Shrivastava (S'12–M'15–SM'19) received his Ph.D. degree from University of Virginia in 2014. Prior to his Ph.D. studies, he worked as a senior design engineer at Texas Instruments, Bangalore from 2006 to 2010. From 2014 to 2016, he worked at the IoT start-up Everactive as senior design director, where he headed the research and development of the energy harvesting and power management solutions. In August 2016, he joined Northeastern University, where he is now working as an Assistant Professor in the Electrical Engineering Department. His research interests include self-powered and ultra-low power circuits and systems, energy-harvesting and analog computing, hardware for AI, internet-of-things, and ultra-low power bio-medical and neural circuits. Dr. Shrivastava is a recipient of 2022 NSF CAREER Award. He currently serves as an Associate Editor for IEEE Open Journal on Circuits and System (OJCAS).