

Article

Nanopower Integrated Gaussian Mixture Model Classifier for Epileptic Seizure Prediction

Vassilis Alimisis *, Georgios Gennis, Konstantinos Touloupas, Christos Dimas , Nikolaos Uzunoglu and Paul P. Sotiriadis

Department of Electrical and Computer Engineering, National Technical University of Athens, 15780 Athens, Greece; giorgosyennis@gmail.com (G.G.); kostastoulou@gmail.com (K.T.); chdim@central.ntua.gr (C.D.); nikolaos.uzunoglu@gmail.com (N.U.); pps@ieee.org (P.P.S.)

* Correspondence: alimisisv@gmail.com

Abstract: This paper presents a new analog front-end classification system that serves as a wake-up engine for digital back-ends, targeting embedded devices for epileptic seizure prediction. Predicting epileptic seizures is of major importance for the patient's quality of life as they can lead to paralyzation or even prove fatal. Existing solutions rely on power hungry embedded digital inference engines that typically consume several μW or even mW . To increase the embedded device's autonomy, a new approach is presented combining an analog feature extractor with an analog Gaussian mixture model-based binary classifier. The proposed classification system provides an initial, power-efficient prediction with high sensitivity to switch on the digital engine for the accurate evaluation. The classifier's circuit is chip-area efficient, operating with minimal power consumption (180 nW) at low supply voltage (0.6 V), allowing long-term continuous operation. Based on a real-world dataset, the proposed system achieves 100% sensitivity to guarantee that all seizures are predicted and good specificity (69%), resulting in significant power reduction of the digital engine and therefore the total system. The proposed classifier was designed and simulated in a TSMC 90 nm CMOS process, using the Cadence IC suite.

Keywords: analog VLSI implementation; analog wake-up; epileptic seizure; Gaussian mixture model; low-power design; seizure prediction



Citation: Alimisis, V.; Gennis, G.; Touloupas, K.; Dimas, C.; Uzunoglu, N.; Sotiriadis, P.P. Nanopower Integrated Gaussian Mixture Model Classifier for Epileptic Seizure Prediction. *Bioengineering* **2022**, *9*, 160. <https://doi.org/10.3390/bioengineering9040160>

Academic Editor: Ali Zarrabi

Received: 22 March 2022

Accepted: 4 April 2022

Published: 5 April 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The continuing progress in integrated circuit (IC) technologies has resulted in complex and power-efficient systems that address the challenges of various Internet of Things (IoT) and machine learning (ML) applications [1,2]. A particular example is wearable systems that monitor the user's health condition, such as electroencephalogram (EEG) monitors [3]. In this case, the subject's brain activity is monitored through the use of electrodes attached to the scalp in order to track, classify, and diagnose epileptic seizures. By continuously monitoring EEG signals throughout the everyday life of the subject, accurate conclusions about their condition can be drawn and intractable epileptic seizures [4,5], which are not amenable to medication, can be forecasted [6,7].

Wearable devices that track the EEG signals can be employed in an everyday fashion. However, the need to operate unobstructed on a Lithium battery or using energy harvesters [8] poses constraints on the acquisition procedure; all-digital signal processing and ML-powered inference can be power-hungry and limit the system's autonomy. A trend to alleviate this limitation is the employment of cascaded classifiers, where the first ones consume relatively low power and are always on, activating more complex units only when needed [9]. Although weight quantization and pruning [10] have notably reduced the power dissipation per inference for digital ML models [11], the digital processing blocks that are provided with the models' input features consume considerable power [12,13].

This renders the aforementioned cascaded classification scheme sub-optimal. To address this, recent work has proposed moving the feature extraction procedure in the analog part of the processing chain [12–15]. Front-end signal processing blocks, like switched capacitor filter banks, operate on the signals prior to the analog-to-digital converter (ADC) to lower the overall system’s power. The digitized features are then input to the ML model in the digital back-end.

To increase the autonomy of wearable EEG monitoring devices, overall power consumption must be decreased below the μW range. Because the energy performance of typical ML models in digital circuitry is in the μW [12,16], an alternative approach seems to be preferable. To this end, in this work we propose an ultra-low power classification system that takes advantage of analog features and uses an analog classifier as a switching device for the power-hungry digital back-end [17]. The proposed architecture, along with the mainstream approaches discussed previously, are illustrated conceptually in Figure 1. The classifier is a Gaussian mixture model (GMM), and its analog implementation consumes 180 nW of power when operating on a 0.6 V supply voltage, in the sub-threshold region. Its predictions are used to switch on and off a subsequent stage of a digital classifier, which provides high accuracy for the whole processing chain. For evaluation, the proposed classifier is designed and verified using a real-world intractable epileptic seizure dataset [4,5].

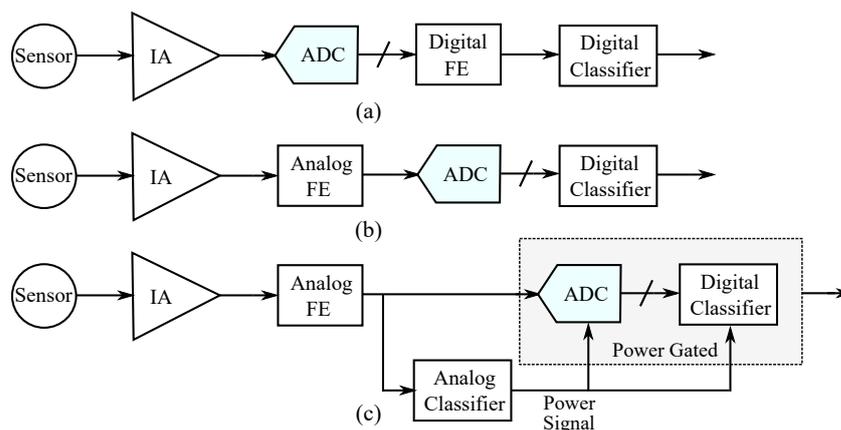


Figure 1. Architecture comparison. (a) All-digital inference. (b) Analog feature extraction block with a digital classifier. (c) Proposed concept architecture, where the digital back-end is turned on and off based on the low power analog classifier’s output.

The remainder of this paper is organized as follows. The background regarding epileptic seizure prediction and analog classifiers is provided in Section 2. Section 3 explains the mathematical foundations of GMMs. The proposed architecture and its building blocks are discussed in Section 4. Section 5 presents the experimental results of the proposed approach on a real-world EEG dataset. A comparison study and discussion are provided in Section 6. Concluding remarks are given in Section 7.

2. Motivation and Background

In this section we provide the necessary background on epileptic seizure prediction and a summary of existing approaches. To introduce the reader to the state-of-the-art in analog implementations of ML systems, a summary of existing analog classifiers is also given.

2.1. Epileptic Seizure Prediction

An epileptic seizure is a sudden excessive neural activity or electrical disturbance in the brain [18,19].

An individual suffering from epilepsy demonstrates symptoms that vary from unnoticeable to paralyzing or even lethal. In practice, patients’ quality of life is severely affected by the unpredictability and the frequency of the seizures. A remedy to this can

be prediction and warning about upcoming epileptic episodes. An accurate prediction of an upcoming seizure could allow them to prepare accordingly and avoid potentially dangerous activities, like, for instance, driving. Epileptic seizure prediction stems from examining the patients’ health using bio-signal acquisition methods.

There are four different states regarding epileptic seizures; (a) pre-ictal, (b) ictal, (c) post-ictal, and (d) inter-ictal [18,19]. States (a)–(c) refer to the periods shortly before, during, and shortly after a seizure, respectively, whereas (d) refers to the period between two seizures, when the patient is considered to be in a normal state. Based on the analysis presented in [20], the duration of the pre-ictal and post-ictal periods varies from 30 min to 2 h. An accurate and real-time identification of the pre-ictal state is crucial, as it is equivalent to predicting an upcoming seizure.

In the literature there are numerous epileptic seizure prediction systems, which follow different approaches to identify the pre-ictal periods. Although the use of EEG signals is the most common approach [20–28], electrocardiograph (ECG) [29,30], electromyograph (EMG) [30], heart rate [30,31], and vibration [31,32] signals have also been used. These systems achieve high accuracy (>80%) on predicting the epileptic seizures. Some of these implementations are edge computing (on-sensor computing) wearable devices [15,21–23], whereas others, to reduce the local power consumption, combine simple data acquisition devices with remote computing [20,24–32] (smartphone or cloud computing). In either case, the low power efficiency of these devices limits their capabilities and usability.

However, there exist multiple architectures that employ analog design methodologies to address epileptic seizure prediction through monitoring EEG signals. The work in [13,15] employs analog feature extraction to greatly minimize the system’s power consumption. A different approach includes employing analog pre-processing circuits directly on the acquisition device in remote computing applications [33–35]. In particular, the analog circuit reduces the overall power consumption of the communication device by reducing the data that need to be transferred to the remote server for prediction. A brief summary of seizure prediction systems in terms of employed algorithms, operating device, and power consumption for all the aforementioned implementations is provided in Table 1.

Table 1. Performance summary for epileptic seizure prediction systems.

Ref.	Model	Device	Power Related Metric
[13]	SVM	hardware	3.07 μ W
[15]	SVM	hardware	66 μ W
[20]	LSTM	software	N/A
[21]	SVM	hardware	7.13 $\frac{\mu\text{J}}{\text{feature}}$
[22]	SVM	hardware	1.35 $\frac{\mu\text{J}}{\text{classification}}$
[23]	Perceptron	hardware	55.89 mW/25 h
[24]	VAE	software	N/A
[25]	CNN	software	N/A
[26]	SVM	software	N/A
[27]	multi-model	smartphone	N/A
[28]	ANN	software	N/A
[29]	custom	(MCU) MSP430 and Cloud	N/A
[30]	custom	smartphone	N/A

Table 1. *Cont.*

Ref.	Model	Device	Power Related Metric
[31]	custom	arduino	N/A
[32]	SDA	cloud	N/A
[33]	filters	hardware	30.4 μ W
[34]	filters	hardware	3.65 mW
[35]	filters	hardware	23 μ W

2.2. Analog Classifiers

Analog integrated circuits (ICs), powered by their capability to operate in the sub-threshold domain [36], are gaining popularity as a means to reduce power consumption in comparison to their digital counterparts. Applications that employ real-time ML techniques are typically power hungry and could greatly benefit from analog circuitry. Nonetheless, analog circuits struggle with high dimensional classification problems as they typically require multiple cascaded multipliers. In practice, analog multipliers are usually unreliable and their operating voltage range is limited. Two main approaches regarding this issue are to either tailor multipliers for specific applications [37,38] or utilize architectures and/or circuits that avoid multipliers [39–43]. Following the former approach, translinear-based, current-mode multipliers [44] are the most popular choice. Regarding the latter, Gaussian function circuits are a commonly used solution [45].

Translinear-based Gaussian function circuits [45] that consist of squaring and exponentiator circuits are utilized in [39,40]. In this case, by leveraging the properties of the exponential function, the multiplication is replaced by the summation of the exponents, which is a trivial task. Alternatively, the work proposed in [41–43] uses more compact building blocks, e.g., bump circuits [41], that implement multivariate Gaussian functions without the use of multipliers. A performance summary of the aforementioned work is presented in Table 2.

Table 2. Performance summary for analog classifiers.

Ref.	Technology	Model	Dimensions	Power Consumption	Area
[37]	0.5 μ m	SVM	14	840.0 nW	9.000 mm ²
[38]	180 nm	LSTM	16 \times 16 matrix	460.3 mW	9.990 mm ²
[39]	180 nm	SVM	64	N/A	0.125 mm ²
[40]	180 nm	K-means	164	N/A	N/A
[41]	90 nm	Bayesian	5	365 nW	0.030 mm ²
[42]	180 nm	SVM	2	220.0 μ W	0.060 mm ²
[43]	0.5 μ m	RBF NN	2	N/A	2.250 mm ²

By examining Table 2, the implementation with the lowest power consumption is [41] (365 nW). This is due to the combination of a compact and simple ML model with ultra low-power building blocks that operate in the sub-threshold domain with a low supply voltage. Based on our previous work in [41], here we build a low power analog classifier and improve upon the accuracy by employing a GMM model instead of a simple Gaussian one.

3. Gaussian Mixture Model

In this section, the mathematical foundations of GMMs, which comprise the core of the proposed classifier, are given. In addition, the use of the GMMs within the scope of classification is also described.

Consider an N -dimensional random variable $X = [x_1, \dots, x_n]$ and its probability density function (PDF) p with $X \sim p$. The GMM is a probabilistic model that consists of a weighted sum of Gaussian distributions and can be used to approximate unknown PDFs from data [46]. In the case of X , the GMM's Gaussian distributions, also noted as components, are also N -dimensional. GMMs belong to the general class of mixture models (MMs) and are widely used in the literature, as they combine both the approximation capabilities of MMs and the properties of Gaussian distributions.

The approximate PDF of X , as modeled by a GMM λ , is given by

$$p(X|\lambda) = \sum_{i=1}^K w_i \cdot \mathcal{N}(X|M_i, \Sigma_i). \tag{1}$$

Here, the component count is $K \geq 1$ and for the weights it holds that $\sum_{i=1}^K w_i = 1$. Each of the components is an N -dimensional Gaussian distribution with a $(N \times 1)$ mean vector M_i and a $(N \times N)$ covariance matrix Σ_i , for $i = 1, \dots, K$.

In the special case of diagonal covariance matrices, each Gaussian distribution is given by

$$\mathcal{N}(X|M_i, \Sigma_i) = \prod_{n=1}^N \mathcal{N}(x_n|\mu_n^i, (\sigma_n^i)^2), \tag{2}$$

where superscript 'i' denotes the Gaussian component and subscript 'n' the dimension, i.e., μ_n^i is the n th component of vector M_i and σ_n^i is the n th component of the diagonal of matrix Σ_i . Hence, each component is derived by the product of N univariate Gaussian distributions given by

$$\mathcal{N}(x_n|\mu_n, (\sigma_n)^2) = \frac{1}{\sqrt{(2\pi) \cdot (\sigma_n)^2}} e^{-\frac{1}{2} \cdot \frac{(x_n - \mu_n)^2}{(\sigma_n)^2}}. \tag{3}$$

GMMs are adapted to data by using the expectation-maximization (EM) algorithm [46]. Although their unsupervised nature renders them suitable for clustering problems, they can also be used within the scope of supervised classification models. Considering a dataset \mathcal{D} with N -dimensional input vectors and C classes, one can fit C separate GMMs $[\lambda_i]_{i=1}^C$ to each subset of \mathcal{D} associated with each class. Therefore, the PDF of the input vectors that belong to each class is approximated by a GMM.

Using the above setting, one can infer the class y of a new input vector X of an unknown class as the one whose approximate PDF provides the highest likelihood, i.e.,

$$\begin{aligned} y &= \operatorname{argmax}_{c \in [1, C]} \left\{ p(X|\lambda_c) \right\} \\ &= \operatorname{argmax}_{c \in [1, C]} \left\{ \sum_{i=1}^K w_i^c \cdot \mathcal{N}(X|M_i^c, \Sigma_i^c) \right\}. \end{aligned} \tag{4}$$

In this case, superscript 'c' denotes the class. It is important to note that in Equation (4) all GMMs share the same number of components K . In the supervised setting, K is denoted as clusters, and this is the naming this paper follows for the rest of the sections. The number of clusters is a hyperparameter of the overall classifier and it is chosen based on the complexity of the data.

4. Proposed Architecture

In this section, the architecture of the proposed analog classifier and the operation of its building blocks are analyzed. To reduce the overall power consumption, in the following building blocks, all transistors operate in the sub-threshold region, and the power supply rails are set to $V_{DD} = -V_{SS} = 0.3V$ for the entire classifier.

Based on Section 3, a GMM-based classifier requires two basic building blocks: one that generates a Gaussian PDF, as in (1), and another that implements the argmax operator, as

in (4). In the case of analog hardware, bump circuits have been proposed for the hardware implementation of a univariate Gaussian PDF [47]. Recently, a modified version of the bump circuit was proposed to generate multivariate PDFs as well [48]. Concerning the analog implementation of the argmax operator, winner-take-all (WTA) circuits have been employed in the literature [49]. In this work, we modify a typical bump circuit and use it in the proposed classifier in order to increase its accuracy.

The modified bump circuit is a combination of two sub-circuits: a symmetric current correlator [48] and a differential block [50]. The aim of this modification is to increase the quality of the Gaussian curve and reduce the distortion in the case of the multivariate bump circuits. In particular, the symmetric current correlator improves the symmetry of the Gaussian curve around the mean value [48]. The simple differential block offers good control of the Gaussian curve’s parameters with a minimal area [50]. The cascode mirrors are used instead of the standard ones, to offer robust mirroring even for small bias currents. This is necessary for multivariate bump circuits. This bump circuit, shown in Figure 2, provides a more accurate Gaussian curve, shown in Figure 3, than either of [48,50]. Transistors’ dimensions are summarized in Table 3. The mean value, the variance, and the height of the Gaussian curve are controlled via the voltage parameters V_r and V_c and the bias current I_{bias} , respectively [48,50].

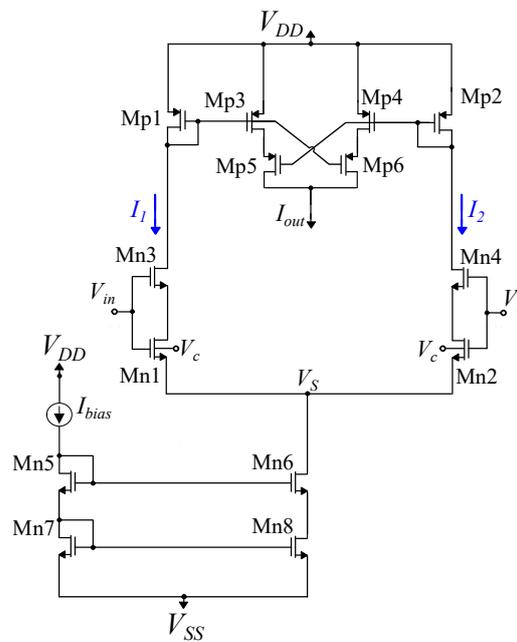


Figure 2. The proposed analog architecture implementing a Gaussian function (bump circuit). V_{in} , V_r , V_c , and I_{bias} are the input voltage, the voltages controlling the mean value, and the variance and the bias current controlling the height of the Gaussian curve, respectively.

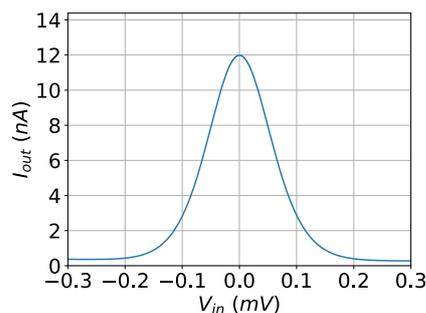


Figure 3. The output of the bump circuit for $I_{bias} = 12$ nA, $V_r = 0$ V, and $V_c = 0$ V.

Table 3. MOS transistors’ dimensions (Figure 2).

Differential Block	W/L ($\mu\text{m}/\mu\text{m}$)	Current Correlator	W/L ($\mu\text{m}/\mu\text{m}$)
M_{n1}, M_{n2}	2.0/1.0	M_{p1}, M_{p2}	0.8/1.6
M_{n3}, M_{n4}	2.0/0.1	$M_{p3}-M_{p6}$	0.4/1.6
$M_{n5}-M_{n7}$	0.4/1.6	-	-
M_{n8}	1.6/1.6	-	-

The multivariate Gaussian PDF is realized by multiplying two or more bump circuits, as described in Equation (2). Consider a sequence of two bump circuits. Biasing the second one with the output current of the first one results in an overall output current that is equivalent to the multiplication of their respective Gaussian curves [48]. An implementation of a 4D Gaussian PDF (four cascaded bump circuits) is shown in Figure 4. Only the first bump is biased with a preset current (I_{bias}), representing the weight w_i of the corresponding cluster i . The topology in Figure 4 constitutes a cluster of the proposed GMM-based classifier.

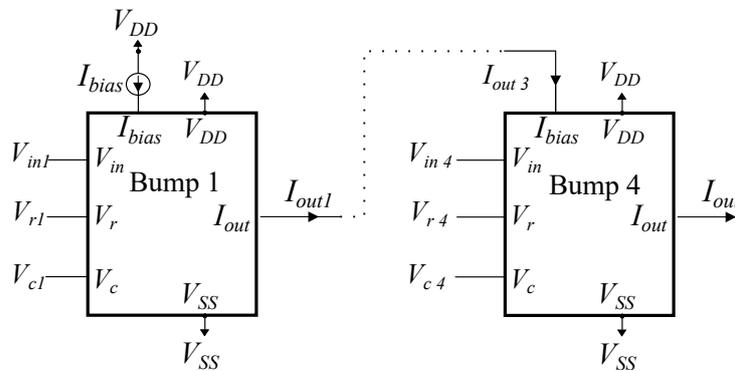


Figure 4. A 4D Bump circuit implementation composed of four sequentially connected univariate bump circuits.

The second block of the proposed architecture is a Lazzaro WTA circuit [49]. Its flexibility and simplicity make it the most popular choice for the implementation of the argmax operator. This WTA circuit is composed of sub-blocks denoted as neuron cells. For a C class classification problem the number of neuron cells must be also C , each one responsible for a single class. In particular, each neuron cell receives the likelihood from a specific GMM and outputs a current in binary format; if the GMM corresponds to the class with the highest likelihood, this current is logical one (which is close to the WTA’s bias current), otherwise it is logical zero (less than 100 pA). For demonstration purposes, a transistor level implementation of a WTA circuit with two neurons is shown in Figure 5. All transistors’ dimensions are equal to $W/L = 0.4 \mu\text{m}/1.6 \mu\text{m}$.

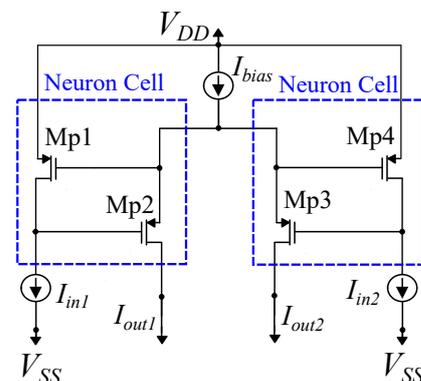


Figure 5. The Lazzaro WTA circuit composed of two PMOS-based neuron circuits.

Utilizing the aforementioned building blocks and based on Equation (4), the proposed GMM-based classifier with two classes, two clusters per class, and 4D inputs is shown in Figure 6. Each GMM class is comprised of two 4D bump circuits, which correspond to the two clusters, and two current mirrors that are used to add the output currents of each cluster. The overall output current $[I_{ci}]_{i=1}^2$ of each class is analogous to the class' likelihood. The WTA circuit compares these probabilities and the predicted class is determined via the currents $[I_i]_{i=1}^2$.

It should be noted that it is impractical to provide the classifier's 34 controlling parameters ($[V_{rj}]_{i=j}^{16}$, $[V_{cj}]_{j=1}^{16}$ and $[I_{biasi}]_{i=1}^2$) externally. Therefore, an alternative option that involves integrating analog memories adjacent to the classifier is preferable. In particular, as typically the classifier will be trained only once prior to its deployment, non-volatile analog memories are a promising choice [51,52]. However, for a general purpose classifier that may require altering this configuration multiple times, dynamic memories can be a more opportune solution [53,54].

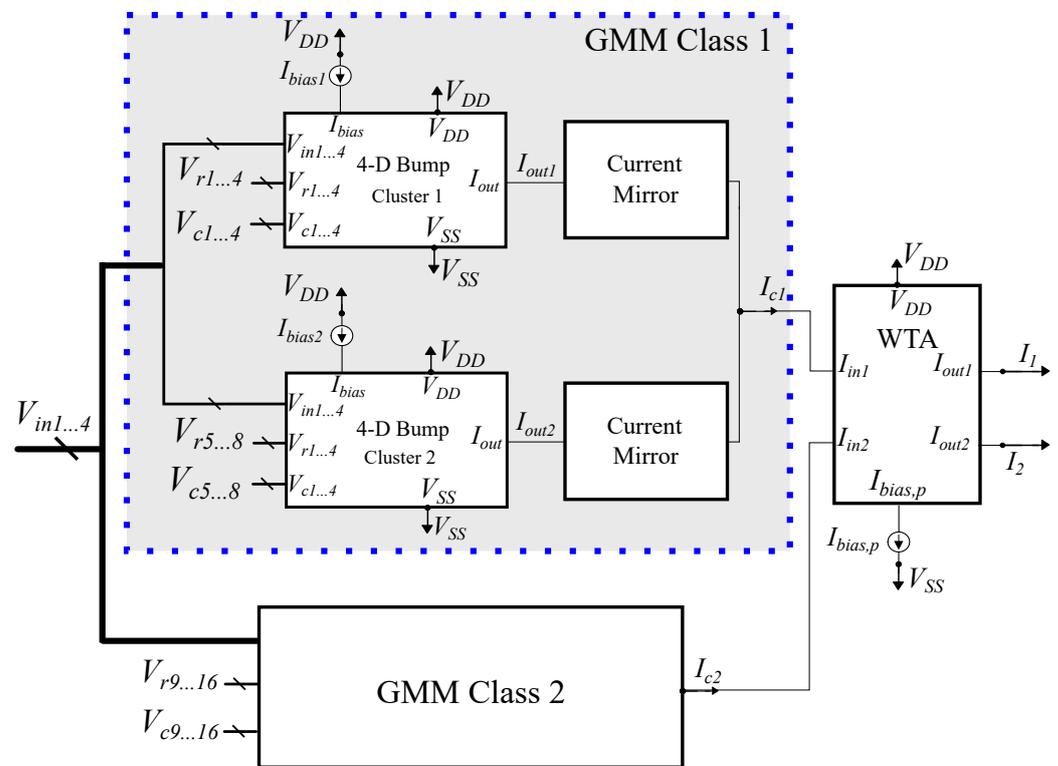


Figure 6. Analog GMM-based classifier with two classes, two clusters per class, and 4D inputs (four input features). (left) GMM class with two clusters; (right) WTA circuit. GMM class 2 follows the same architecture as the depicted GMM class 1.

5. Epileptic Seizure Prediction Application

In this section, the proposed classifier is tested on a real-world epilepsy seizure prediction problem [4,5] to confirm its proper operation. The classifier has been designed using the Cadence IC suite in a TSMC 90 nm CMOS process. All simulation results are conducted on the layout (post-layout simulations), which is shown in Figure 7.

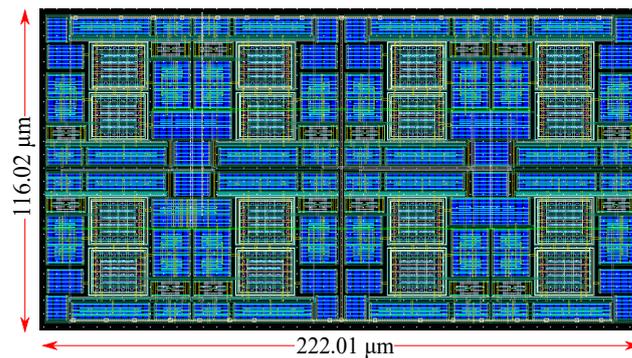


Figure 7. Layout of the proposed classifier circuit.

The data are acquired from the CHB-MIT Scalp EEG database [4,5] and contain EEG signals from children with intractable epilepsy. The ictal periods are labeled by expert physicians. Here, pre-ictal and post-ictal periods are considered to span an hour before and an hour after the seizure, respectively. The data samples that do not belong in ictal, pre-ictal, or post-ictal periods are labeled as inter-ictal.

There are four features for the classification: the signal's peak-to-peak voltage and energy percentages in the alpha and the first and second half of the gamma frequency bands [55]. These features can be efficiently derived from the raw EEG signals using analog feature extraction techniques [13,56]. The system's necessary parameters are derived by software-based training, prior to the circuit's deployment.

The aim of the classifier is to successfully distinguish the pre-ictal from the inter-ictal periods. In order to operate as a minimal power front-end wake-up circuit, it must predict all possible seizures and maintain a low number of false positive alarms. The first requirement is equivalent to having high classification sensitivity [57], which is measured by:

$$\text{sensitivity} = \frac{\text{Predicted Seizures}}{\text{Predicted Seizures} + \text{Missed Seizures}}. \quad (5)$$

Achieving a high sensitivity score is crucial for the patient's health, as it ensures that all upcoming seizures will be predicted. However, the second requirement is equivalent to minimizing the rate with which the high power consumption digital back-end is turned on. This leads to a significant power consumption reduction for the whole system, shown in Figure 1c. An appropriate measure to quantify this reduction is the specificity [57] of the analog classifier, given by:

$$\text{specificity} = \frac{\text{True Negative}}{\text{True Negative} + \text{False Positive}}. \quad (6)$$

In practice, this metric is the ratio of the time that the digital back-end is idle to the duration of all the inter-ictal periods (no risk for seizure).

To test the proposed classifier both in terms of classification specificity and circuit's behavior in PVT variations, two separate tests are conducted. The first one is a comparison between the proposed implementation and a software-based one. In particular, 20 separate software-based training iterations are conducted to account for random effects. The resulting specificity scores are summarized in Table 4. The proposed architecture's mean specificity is only 2% lower than that of a software-based implementation. For demonstration purposes, the state of four patients along with the predictions of the analog classifier are presented in Figure 8. The classifier successfully predicts all 17 seizures (100% sensitivity) of the test set. The second test is a Monte-Carlo analysis for $N = 100$ points, for one of the previous 20 candidates. The Monte-Carlo analysis histogram is shown in Figure 9. Its mean value is $\mu_M = 69.93\%$ with a standard deviation of $\sigma_M = 0.41\%$. This confirms the proper performance and operation of the proposed architecture.

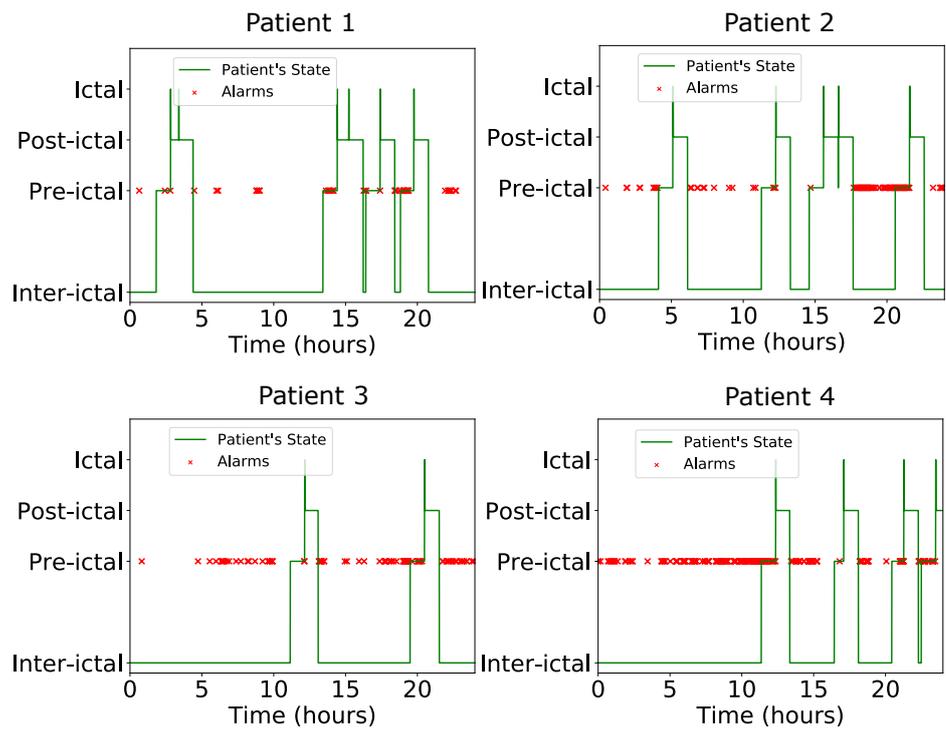


Figure 8. The alarms triggered by the classifier for four patients in a 24-h period. The ideal behavior is the raising of at least one alarm in each pre-ictal period, without raising alarms during the inter-ictal ones. The ictal and post-ictal regions are irrelevant for the classifier.

Table 4. Specificity (over 20 iterations).

Method	Best	Worst	Mean	Std.
Software	71.30%	71.08%	71.27%	0.07%
Proposed	70.65%	67.39%	69.07%	0.51%

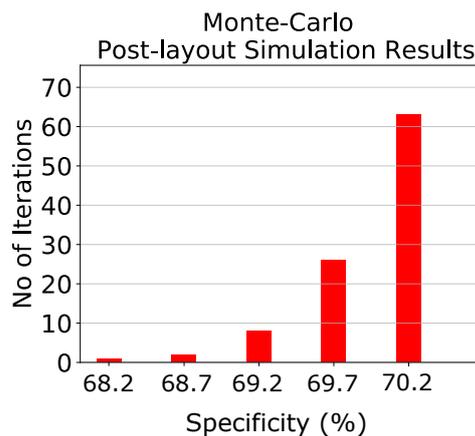


Figure 9. Post-layout Monte-Carlo sensitivity analysis simulation results on the specificity of the classifier for one of the previous 20 iterations.

6. Discussion and Comparison

A comparison between this work and other studies that employ analog design methodologies to address epileptic seizure prediction through monitoring EEG signals is provided in Table 5. Here, it is seen that this work achieves very low power consumption per channel (180 nW per channel), outperforming all the implementations except that from [13], which achieves 96 nW per channel. Nonetheless, as the proposed implementation requires only

a single channel, its total power consumption is significantly smaller. In particular, the proposed architecture consumes power in the range of nW, which is not the case for the rest of the implementations in Table 5. This power dissipation is achieved using a supply voltage of only 0.6 V, which is also the lowest one in Table 5. The specificity of the proposed classifier is 69%, which, along with [13] (86%) and [15] (84.4%), constitutes the three highest specificity scores.

Another important metric for measuring efficiency in analog computing, which is invariant to the application and is therefore a relatively fair metric for comparing architectures designed for different applications, is the energy consumed per operation. The proposed classifier consumes 180 nW and can achieve a computational speed of 166 K classifications per second, which results in 1.1 pJ per classification. Each classification, for a GMM-based classifier composed of two classes, two clusters per class, and 4D inputs, requires 131 operations. This results in the classifier's consumption being 8.2 fJ per operation. Unfortunately, these metrics are not provided in the literature for comparison purposes.

Table 5. Performance summary for analog epileptic seizure prediction systems.

Ref.	Technology	Power Supply	Power Consumption per Channel	Total Power Consumption	No. of Channels	Specificity
This work	90 nm	0.6 V	180 nW	180 nW	1	69%
[13]	65 nm	N/A	96 nW	3.07 μ W	23	86%
[15]	180 nm	1.8 V	8.25 μ W	66 μ W	8	84.4%
[33]	350 nm	1.25 V	950 nW	30.4 μ W	32	55%
[34]	90 nm	1.25 V	1.14 μ W	3.65 mW	32	48.5%
[35]	180 nm	1.8V	23 μ W	23 μ W	1	50%

As shown in Table 5, most epileptic seizure prediction systems employ multiple channels, i.e., electrodes, in order to increase their accuracy. Nonetheless, acquisition devices with multiple electrodes are usually uncomfortable for the patient and impractical for constant monitoring. To this end, this work focuses on extracting data from a single electrode. By doing so, the resulting device is less bulky and more convenient to use. In addition, to further increase the device's portability, the classifier is proposed to operate in an embedded device. In this way, the patient can be monitored constantly with no requirements for wireless communication with other devices as proposed in [33–35].

In real-world scenarios, EEG signal acquisition is affected by uncontrolled parameters and environmental factors. In the case of a single electrode in particular, motion artifacts, electrode misplacement, and external electromagnetic interference can drastically reduce the quality of the signal and potentially lead to diagnostic errors. Having multiple electrodes for signal acquisition may seem more robust, as the contaminated recordings could be only a fraction of the total inputs to the prediction system, but this comes at the cost of the system's portability and with no theoretical guarantee. Efforts to determine the goodness of the acquired signals have been proposed in the literature via employing ML classification techniques [58]. We argue that this quality assessment can implicitly take place within the GMM classifier of our system provided that: (a) real-world, noise-contaminated EEG signals are used for training and, (b) the classifier is expanded to provide confidence bounds about its predictions. By doing so, additional systems for quality assessment, as in [58], can be avoided and thereby the area and power consumption of the device is unchanged.

Another important design consideration is the trade-off between the wake-up circuit's power consumption and its specificity. As the specificity of the wake-up circuit increases, the overall power consumption of the digital circuit, which is typically greater than the analog one's, decreases. However, to achieve high specificity values, it is essential to increase the complexity of the analog circuit. In particular, to improve the classifier's performance, improved-performance acquisition devices, more analog feature extraction

circuits, and larger analog memories storing the classifier's parameters are required. All the aforementioned modifications result in increased power consumption. In practice, increasing the power consumption of the analog front-end must be done cautiously; a classification system with a power greedy analog classifier that switches on and off a digital one may consume more power than an all-digital one.

7. Conclusions

A fully analog processing unit was presented as an alternative to the conventional front-end architectures for inference systems targeting EEG signals. The proposed system includes an 180 nW or 8.2 fJ per operation analog integrated GMM-based classifier, which activates the high-performance digital inference back-end only when needed. Its main building blocks are Gaussian function circuits and the Lazzaro WTA circuit. The classifier was trained on a real-world seizure prediction dataset and designed in a TSMC 90 nm technology. Post-layout simulation results suggest that the proposed circuit achieves 100% sensitivity, as all 17 seizures of the test set are predicted, and 69.07% specificity.

Author Contributions: Investigation, V.A., G.G., K.T. and C.D.; writing—original draft, V.A., G.G., K.T. and C.D.; writing—review and editing, V.A., G.G., K.T., C.D., N.U. and P.P.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research is co-financed by Greece and the European Union (European Social Fund- ESF) through the Operational Programme "Human +Resources Development, Education and Lifelong Learning" in the context of the project "Strengthening Human Resources Research Potential via Doctorate Research" (MIS-5000432), implemented by the State Scholarships Foundation (IKY).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used in this study are openly available in CHB-MIT Scalp EEG Database at <https://physionet.org/content/chbmit/1.0.0/> (accessed on 21 March 2022), reference number [4].

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ishtiaq, A.; Khan, M.U.; Ali, S.Z.; Habib, K.; Samer, S.; Hafeez, E. A Review of System on Chip (SOC) Applications in Internet of Things (IOT) and Medical. In Proceedings of the International Conference on Advances in Mechanical Engineering, ICAME21, Islamabad, Pakistan, 27–28 January 2021; pp. 1–10.
2. Tsai, C.H.; Yu, W.J.; Wong, W.H.; Lee, C.Y. A 41.3/26.7 pJ per neuron weight RBM processor supporting on-chip learning/inference for IoT applications. *IEEE J. Solid-State Circuits* **2017**, *52*, 2601–2612. [CrossRef]
3. Casson, A.J.; Yates, D.C.; Smith, S.J.; Duncan, J.S.; Rodriguez-Villegas, E. Wearable electroencephalography. *IEEE Eng. Med. Biol. Mag.* **2010**, *29*, 44–56. [CrossRef] [PubMed]
4. CHB-MIT Scalp EEG Database. Available online: <https://physionet.org/content/chbmit/1.0.0/> (accessed on 21 March 2022).
5. Goldberger, A.L.; Amaral, L.A.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, 215–220. [CrossRef] [PubMed]
6. Subasi, A.; Kevric, J.; Abdullah Canbaz, M. Epileptic seizure detection using hybrid machine learning methods. *Neural Comput. Appl.* **2019**, *31*, 317–325. [CrossRef]
7. Gómez, C.; Arbeláez, P.; Navarrete, M.; Alvarado-Rojas, C.; Le Van Quyen, M.; Valderrama, M. Automatic seizure detection based on imaged-EEG signals through fully convolutional networks. *Sci. Rep.* **2020**, *10*, 21833. [CrossRef]
8. Priya, S.; Inman, D.J. *Energy Harvesting Technologies*; Springer: New York, NY, USA, 2009; Volume 21, p. 2.
9. Goetschalckx, K.; Moons, B.; Lauwereins, S.; Andraud, M.; Verhelst, M. Optimized hierarchical cascaded processing. *IEEE J. Emerg. Sel. Top. Circuits Syst.* **2020**, *8*, 884–894. [CrossRef]
10. Hubara, I.; Courbariaux, M.; Soudry, D.; El-Yaniv, R.; Bengio, Y. Quantized neural networks: Training neural networks with low precision weights and activations. *J. Mach. Learn. Res.* **2017**, *18*, 6869–6898.
11. Zhang, K.; Ying, H.; Dai, H.N.; Li, L.; Peng, Y.; Guo, K.; Yu, H. Compacting Deep Neural Networks for Internet of Things: Methods and Applications. *IEEE Internet Things J.* **2021**, *8*, 11935–11959. [CrossRef]
12. Villamizar, D.A.; Muratore, D.G.; Wieser, J.B.; Murmann, B. An 800 nW Switched-Capacitor Feature Extraction Filterbank for Sound Classification. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2021**, *68*, 1578–1588. [CrossRef]

13. Zhang, Y.; Mirchandani, N.; Onabajo, M.; Shrivastava, A. RSSI Amplifier Design for a Feature Extraction Technique to Detect Seizures with Analog Computing. In Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, 12–14 October 2020; pp. 1–5.
14. Yang, M.; Liu, H.; Shan, W.; Zhang, J.; Kiselev, I.; Kim, S.J.; Enz, C.; Seok, M. Nanowatt acoustic inference sensing exploiting nonlinear analog feature extraction. *IEEE J. Solid-State Circuits* **2021**, *56*, 3123–3133. [[CrossRef](#)]
15. Yoo, J.; Yan, L.; El-Damak, D.; Altaf, M.A.B.; Shoeb, A.H.; Chandrakasan, A.P. An 8-channel scalable EEG acquisition SoC with patient-specific seizure classification and recording processor. *IEEE J. Solid-State Circuits* **2012**, *48*, 214–228. [[CrossRef](#)]
16. De Vita, A.; Pau, D.; Parrella, C.; Di Benedetto, L.; Rubino, A.; Licciardo, G.D. Low-power HWAccelerator for AI edge-computing in human activity recognition systems. In Proceedings of the 2020 2nd IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS), Genova, Italy, 31 August–2 September 2020; pp. 291–295.
17. Yip, M.; Bohorquez, J. L.; Chandrakasan, A. P. A 0.6 V 2.9 μ W mixed-signal front-end for ECG monitoring. In Proceedings of the 2012 Symposium on VLSI Circuits (VLSIC), Honolulu, HI, USA, 13–15 June 2012; pp. 66–67.
18. Karoly, P.J.; Rao, V.R.; Gregg, N.M.; Worrell, G.A.; Bernard, C.; Cook, M.J.; Baud, M.O. Cycles in epilepsy. *Nat. Rev. Neurol.* **2021**, *17*, 267–284. [[CrossRef](#)] [[PubMed](#)]
19. World Health Organization; Global Campaign against Epilepsy; Programme for Neurological Diseases; Neuroscience (World Health Organization); International Bureau for Epilepsy; World Health Organization, Department of Mental Health; International League against Epilepsy. *Atlas: Epilepsy Care in the World*; World Health Organization: Geneva, Switzerland, 2005.
20. Tsiouris, K.M.; Pezoulas, V.C.; Zervakis, M.; Konitsiotis, S.; Koutsouris, D.D.; Fotiadis, D.I. A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. *Comput. Biol. Med.* **2018**, *99*, 24–37. [[CrossRef](#)] [[PubMed](#)]
21. Shoaib, M.; Jha, N.K.; Verma, N. A compressed-domain processor for seizure detection to simultaneously reduce computation and communication energy. In Proceedings of the IEEE 2012 Custom Integrated Circuits Conference, San Jose, CA, USA, 9–12 September 2012; pp. 1–4.
22. Zhang, J.; Huang, L.; Wang, Z.; Verma, N. A seizure-detection IC employing machine learning to overcome data-conversion and analog-processing non-idealities. In Proceedings of the 2015 IEEE Custom Integrated Circuits Conference (CICC), San Jose, CA, USA, 28–30 September 2015; pp. 1–4.
23. Lin, S.K.; Wang, L.C.; Lin, C.Y.; Chiueh, H. An ultra-low power smart headband for real-time epileptic seizure detection. *IEEE J. Transl. Eng. Health Med.* **2018**, *6*, 1–10. [[CrossRef](#)]
24. Abdelhameed, A.M.; Bayoumi, M. An Efficient Deep Learning System for Epileptic Seizure Prediction. In Proceedings of the 2021 IEEE International Symposium on Circuits and Systems (ISCAS), Daegu, Korea, 22–28 May 2021; pp. 1–5.
25. O’Shea, A.; Lightbody, G.; Boylan, G.; Temko, A. Neonatal seizure detection from raw multi-channel EEG using a fully convolutional architecture. *Neural Netw.* **2020**, *123*, 12–25. [[CrossRef](#)]
26. Shoeb, A.H.; Gutttag, J.V. Application of machine learning to epileptic seizure detection. In Proceedings of the ICML, Haifa, Israel, 21–24 June 2010.
27. Lasefr, Z.; Reddy, R.R.; Elleithy, K. Smart phone application development for monitoring epilepsy seizure detection based on EEG signal classification. In Proceedings of the 2017 IEEE 8th Annual Ubiquitous Computing, Electronics and Mobile Communication Conference (UEMCON), New York, NY, USA, 19–21 October 2017; pp. 83–87.
28. Tzallas, A.T.; Tsipouras, M.G.; Fotiadis, D.I. Epileptic seizure detection in EEGs using time–frequency analysis. *IEEE Trans. Inf. Technol. Biomed.* **2009**, *13*, 703–710. [[CrossRef](#)]
29. Chen, H.; Gu, X.; Mei, Z.; Xu, K.; Yan, K.; Lu, C.; Wang, L.; Shu, F.; Xu, Q.; Oetomo, S.B.; et al. A wearable sensor system for neonatal seizure monitoring. In Proceedings of the 2017 IEEE 14th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Eindhoven, The Netherlands, 9–12 May 2017; pp. 27–30.
30. Gheryani, M.; Salem, O.; Mehaoua, A. An effective approach for epileptic seizures detection from multi-sensors integrated in an Armband. In Proceedings of the 2017 IEEE 19th International Conference on e-Health Networking, Applications and Services (Healthcom), Dalian, China, 12–15 October 2017; pp. 1–6.
31. Ramirez-Alaminos, J.M.; Sendra, S.; Lloret, J.; Navarro-Ortiz, J. Low-cost wearable bluetooth sensor for epileptic episodes detection. In Proceedings of the 2017 IEEE International Conference on Communications (ICC), Paris, France, 21–25 May 2017; pp. 1–6.
32. Marquez, A.; Dunn, M.; Ciriaco, J.; Farahmand, F. iSeiz: A low-cost real-time seizure detection system utilizing cloud computing. In Proceedings of the 2017 IEEE Global Humanitarian Technology Conference (GHTC), San Jose, CA, USA, 19–22 October 2017; pp. 1–7.
33. Iranmanesh, S.; Rodriguez-Villegas, E. A 950 nW analog-based data reduction chip for wearable EEG systems in epilepsy. *IEEE J. Solid-State Circuits* **2017**, *52*, 2362–2373. [[CrossRef](#)]
34. Iranmanesh, S.; Raikos, G.; Imtiaz, S.A.; Rodriguez-Villegas, E. A seizure-based power reduction SoC for wearable EEG in epilepsy. *IEEE Access* **2019**, *7*, 151682–151691. [[CrossRef](#)]
35. Imtiaz, S.A.; Iranmanesh, S.; Rodriguez-Villegas, E. A low power system with EEG data reduction for long-term epileptic seizures monitoring. *IEEE Access* **2019**, *7*, 71195–71208. [[CrossRef](#)]
36. Liu, S.C.; Kramer, J.; Indiveri, G.; Delbrück, T.; Douglas, R. *Analog VLSI: Circuits and Principles*; MIT Press: Cambridge, MA, USA, 2002.

37. Chakrabartty, S.; Cauwenberghs, G. Sub-microwatt analog VLSI trainable pattern classifier. *IEEE J. Solid-State Circuits* **2007**, *42*, 1169–1179. [[CrossRef](#)]
38. Zhao, Z.; Srivastava, A.; Peng, L.; Chen, Q. Long short-term memory network design for analog computing. *ACM J. Emerg. Technol. Comput. Syst. (JETC)* **2019**, *15*, 1–27. [[CrossRef](#)]
39. Zhang, R.; Shibata, T. Fully parallel self-learning analog support vector machine employing compact gaussian generation circuits. *Jpn. J. Appl. Phys.* **2012**, *51*, 04DE10. [[CrossRef](#)]
40. Zhang, R.; Shibata, T. An analog on-line-learning K-means processor employing fully parallel self-converging circuitry. *Analog Integr. Circuits Signal Process.* **2013**, *75*, 267–277. [[CrossRef](#)]
41. Alimisis, V.; Gennis, G.; Dimas, C.; Sotiriadis, P.P. An Analog Bayesian Classifier Implementation, for Thyroid Disease Detection, based on a Low-Power, Current-Mode Gaussian Function Circuit. In Proceedings of the 2021 International Conference on Microelectronics (ICM), New Cairo City, Egypt, 19–22 December 2021; pp. 153–156.
42. Kang, K.; Shibata, T. An on-chip-trainable Gaussian-kernel analog support vector machine. *IEEE Trans. Circuits Syst. I Regul. Pap.* **2009**, *57*, 1513–1524. [[CrossRef](#)]
43. Peng, S.Y.; Hasler, P.E.; Anderson, D. An analog programmable multi-dimensional radial basis function based classifier. In Proceedings of the 2007 IFIP International Conference on Very Large Scale Integration, Atlanta, GA, USA, 15–17 October 2007; pp. 13–18.
44. Lopez-Martin, A.J.; Carlosena, A. Current-mode multiplier/divider circuits based on the MOS translinear principle. *Analog Integr. Circuits Signal Process.* **2001**, *28*, 265–278. [[CrossRef](#)]
45. Alimisis, V.; Gourdouparis, M.; Gennis, G.; Dimas, C.; Sotiriadis, P.P. Analog Gaussian Function Circuit: Architectures, Operating Principles and Applications. *Electronics* **2021**, *10*, 2530. [[CrossRef](#)]
46. Bishop, C.M.; Nasrabadi, N.M. *Pattern Recognition and Machine Learning*; Springer: New York, NY, USA, 2006.
47. Delbrueck, T.; Mead, C. Bump circuits. In Proceedings of the International Joint Conference on Neural Networks, Nagoya, Japan, 25–29 October 1993; Volume 1, pp. 475–479.
48. Alimisis, V.; Gourdouparis, M.; Dimas, C.; Sotiriadis, P.P. A 0.6 V, 3.3 nW, Adjustable Gaussian Circuit for Tunable Kernel Functions. In Proceedings of the 2021 34th SBC/SBMicro/IEEE/ACM Symposium on Integrated Circuits and Systems Design (SBCCI), Campinas, Brazil, 23–27 August 2021; pp. 1–6.
49. Lazzaro, J.; Ryckebusch, S.; Mahowald, M.A.; Mead, C.A. Winner-take-all networks of O (n) complexity. In Proceedings of the Advances in Neural Information Processing Systems 1 (NIPS 1988), Denver, CO, USA, 27–30 November 1988; Volume 1.
50. Alimisis, V.; Gourdouparis, M.; Dimas, C.; Sotiriadis, P.P. Ultra-Low Power, Low-Voltage, Fully-Tunable, Bulk-Controlled Bump Circuit. In Proceedings of the 2021 10th International Conference on Modern Circuits and Systems Technologies (MOCAST), Thessaloniki, Greece, 5–7 July 2021; pp. 1–4.
51. Chakrabartty, S.; Cauwenberghs, G. Fixed-current method for programming large floating-gate arrays. In Proceedings of the 2005 IEEE International Symposium on Circuits and Systems, Kobe, Japan, 23–26 May 2005; pp. 3934–3937.
52. Chakrabartty, S.; Cauwenberghs, G. Sub-microwatt analog VLSI support vector machine for pattern classification and sequence estimation. In Proceedings of the Advances in Neural Information Processing Systems 17 (NIPS 2004), Vancouver, BC, Canada, 13–18 December 2004; Volume 17.
53. Cauwenberghs, G.; Yariv, A. Fault-tolerant dynamic multilevel storage in analog VLSI. *IEEE Trans. Circuits Syst. II Analog Digital Signal Process.* **1994**, *41*, 827–829. [[CrossRef](#)]
54. Hock, M.; Hartel, A.; Schemmel, J.; Meier, K. An analog dynamic memory array for neuromorphic hardware. In Proceedings of the 2013 European Conference on Circuit Theory and Design (ECCTD), Dresden, Germany, 8–12 September 2013; pp. 1–4.
55. Miller, R. Theory of the normal waking EEG: from single neurones to waveforms in the alpha, beta and gamma frequency ranges. *Int. J. Psychophysiol.* **2007**, *64*, 18–23. [[CrossRef](#)]
56. Chen, M.; Boric-Lubecke, O.; Lubecke, V.M. 0.5- μm CMOS Implementation of Analog Heart-Rate Extraction With a Robust Peak Detector. *IEEE Trans. Instrum. Meas.* **2008**, *57*, 690–698. [[CrossRef](#)]
57. Altman, D.G.; Bland, J.M. Diagnostic tests. 1: Sensitivity and specificity. *BMJ Br. Med. J.* **1994**, *308*, 1552. [[CrossRef](#)]
58. Gosselin, F.; Navarro-Sune, X.; Vozzi, A.; Pandremmenou, K.; De Vico Fallani, F.; Attal, Y.; Chavez, M. Quality assessment of single-channel EEG for wearable devices. *Sensors* **2019**, *19*, 601. [[CrossRef](#)] [[PubMed](#)]