This article has been accepted for publication in IEEE Transactions on Biomedical Circuits and Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBCAS.2024.3396115

An Efficient Brain-Switch for Asynchronous Brain-Computer Interfaces

Daniel Valencia[1,2], Patrick P. Mercier[2], and Amir Alimohammad[1]
Department of Electrical and Computer Engineering
[1]San Diego State University, San Diego, U.S.A.
[2]University of California San Diego, La Jolla, U.S.A.

***Abstract*—Intracortical brain computer interfaces (iBCIs) utilizing extracellular recordings mainly employ in vivo signal processing application-specific integrated circuits (ASICs) to detect action potentials (spikes). Conventionally, "brain-switches" based on spiking activity have been employed to realize asynchronous (self-paced) iBCIs, estimating when the user involves in the underlying BCI task. Several studies have demonstrated that local field potentials (LFPs) can effectively replace action potentials, drastically reducing the power consumption and processing requirements of in vivo ASICs. This article presents the first LFP-based brain-switch design and implementation using gated recurrent neural networks (RNNs). Compared to the previously reported brain-switches, our design requires no exhaustive learning phase for the estimation of optimal recording channels or frequency band selection, making it more applicable to practical asynchronous iBCIs. The synthesized ASIC of the designed in vivo LFP-based feature extraction unit, in a standard 180-nm CMOS process, occupies only 0.09 mm^2 of silicon area, and the post place-and-route synthesis results indicate that it consumes 91.87 nW of power while operating at 2 kHz. Compared to the previously published ASICs, the proposed LFP-based brain-switch consumes the least power for in vivo digital signal processing and achieves comparable state estimation performance to that of spike-based brain-switches.**

I. Introduction

Over the past decade, researchers have investigated the activities of individual neurons concerning their neighboring neurons and their response to various stimuli [1], [2]. Neurons communicate by firing electrical pulses, called action potentials (APs) or spikes. The electrical activity of neurons can be measured and recorded by multi-electrode arrays (MEAs). Each intracortical electrode implanted in the motor cortex records extracellular electrical activity from a relatively small population of neurons within a few hundred micrometers of the neuron closest to the tip of the electrode [3]. This measurement is often contaminated with ambient noise and technical artifacts, such as instrumentation noise. The measured electrical activity inside the gray matter of the brain can be used for muscle control, sensory perception (such as seeing and hearing), as well as speech, decision making, and self-control.

The intracortical brain-computer interfaces (iBCIs) predominantly quantify the neural activity of the brain using multi-unit activities (MUAs). Neural spikes embedded in the MUAs are detected and separated from the background noise by comparing the recorded and filtered voltage waveforms with a threshold, commonly considered as the scaled value of the background noise. Neighboring neurons often fire spikes of similar shape and amplitude; however, relative to their distances to an electrode's tip, the shape of spike waveforms may differ among neurons. This fact allows the spiking activity of individual neurons to be separated through the spike sorting process [4]. It has already been demonstrated that robust iBCIs can be implemented without employing computationally-daunting spike sorting [5]–[9]. In this case, all threshold crossings (TCs) of the recorded and filtered voltage waveforms associated with an electrode (channel) are treated as spikes from one putative neuron. Instead of transmitting spike waveforms for offline neural decoding and control, transmitting the number of spikes over a given time window as the feature of interest would dramatically reduce the transmission data rate and power consumption [10].

Although spike-based decoding is commonly employed, various studies have shown that LFP signals can alternatively be used for neural decoding [11]–[13]. In LFP-based decoding studies, the spectral power is the most commonly employed metric to quantify the neural activity. This is predicated on historical findings indicating a correlation between changes in the neural signal band power and different physiological states, such as oscillations in the gamma band (30 – 90 Hz) during sleep in humans [14], [15]. Similarly, the power in the beta band (15 – 35 Hz) has been linked to movement preparation [16].

Depending on the timing and nature of the interaction between the user's brain activity and the iBCI system, two paradigms exist. In synchronous iBCIs, the user's neural activity is processed within a predefined schedule, such as when the user is instructed to perform the underlying BCI task, such as motor imagery or concentration on a specific stimulus.On the other hand, asynchronous iBCIs allow the user to initiate actions at their own pace without being restricted to specific time intervals or cues. Asynchronous iBCIs aim to offer a more natural and flexible brain-computer interaction.However, to support a self-paced iBCI, the system needs to continuously monitor the user's brain activity. Compared to synchronous BCIs, asynchronous iBCIs are better suited for applications where users require continuous control, such as prosthetic control.

Over the last decade, synchronous iBCIs and the decoding of information, such as movement kinematics (i.e., direction, velocity, etc.), from recorded extracellular signals have received considerable attention. However, the efficient realization of the asynchronous iBCIs and addressing the need to detect when a user desires to use a prosthetic or assistive device have remained relatively understudied. A two-state "active/inactive" "brain-switch" [17]–[20] continuously monitors neural signal features to detect whether the user would like to engage in the underlying BCI activity. Consequently, it only processes neural signals during the "active" state. Another important aspect of brain state estimation is that

This article has been accepted for publication in IEEE Transactions on Biomedical Circuits and Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBCAS.2024.3396115

the inclusion of periods of unintended iBCI control could introduce errors to the decoder analysis. For instance, decoding neural activity between periods of active BCI operations may result in unintended, and potentially dangerous, prosthetic movements when the user is not intended to be engaged in the iBCI operation. To eliminate unnecessary and unintended decoding of neural signals during periods of non-BCI activity, we designed an intention estimation system to discriminate between periods of active BCI control and inactive states. We utilized a recurrent neural network and LFPs as the input feature of interest to detect discrete state changes when the user desires to engage in the underlying BCI activity.

The rest of this article is organized as follows. Section II briefly reviews the motivation for employing LFPs over other signal modalities. Section III discusses the application of LFPs for realizing brain-switches. The relative performance of the designed LFP-based brain-switch is compared against those of the spiking-based models. Section IV presents the hardware architecture and realization of the proposed LFP-based feature extraction unit. Finally, Section V makes some concluding remarks.

II. Feasibility of Employing Local Field Potentials

One invasive method for recording the brain's neural activity is to use penetrating microelectrodes implanted in the related brain area, such as the primary motor cortex. These electrodes provide valuable information for decoding intended muscle movements. Intracortical MEAs, capable of recording hundreds to thousands of channels simultaneously, offer the highest degrees of freedom among neural signal modalities. They have been conventionally used in iBCIs to record and translate neural activities for controlling robotic prostheses and assistive devices. Whether using a single electrode or an MEA, recording action potentials (spikes) from an individual neuron requires placing the electrode tip within 50–100 mm from the neuron. The electrical spike signal best reflects the neural activity of the brain and has a high signal-to-noise ratio (SNR), as well as high temporal and spatial resolutions. Therefore, the firing pattern of spikes is commonly utilized to extract motor-related information effectively.

Depending on the cortical area and layer, however, each electrode may record action potentials of several individual neurons surrounding the electrode tip. If these multiple recorded neurons have distinctive spike shapes, their individual firing rates (i.e., the number of spikes within 30 to 100 ms time bins) can be discriminated by the spike sorting process, yielding single-unit responses, also known as single-unit activity (SUA). If their action potential voltage traces are similar, spike sorting may not be able to differentiate the individual spike waveforms, thus yielding multi-unit activity (MUA), which represent the total firing rates of two or more neurons. MUAs can be obtained by high-pass filtering the raw data at 250 Hz to 5 kHz and counting all waveforms crossing a threshold. An alternative neural signal that has been shown to be well-correlated with MUAs is the spiking band power (SBP), which has a relatively low sampling requirement of 2 kHz [21] and can be considered as a low-frequency version of MUAs.

Fig. 1 illustrates the relative spatial resolution of the alternative signal modalities available for intracortical recording, with the center representing the location of the electrode tip. SUAs have a spatial resolution of single units, on the order of 50 μm, while MUAs have a spatial resolution over 100 μm. The SBP is similar to MUAs but can detect the activity of neurons over a larger area of 250 μm. Additionally, the LFP comprises the activity of neurons within 500 μm [21]–[24].

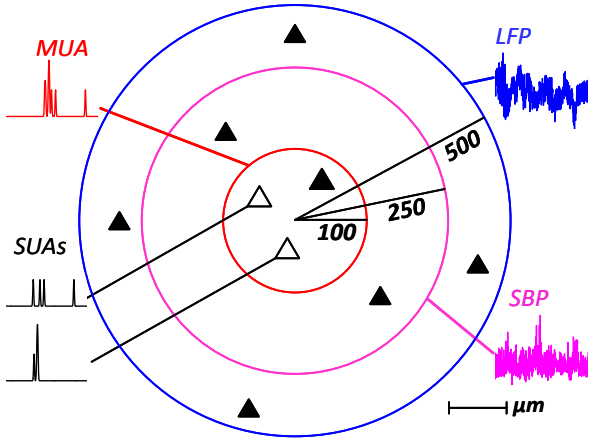

Fig. 1: The different signal modalities of intracortical recording electrodes.

Maintaining signal stability poses a serious challenge for long-term iBCI recordings. One major drawback of spike-based iBCIs is that over time the electrode tips become increasingly encapsulated by brain tissue scarring, preventing the long-term observation of individual neurons [25]. Fortunately, this encapsulation does not significantly affect lower frequency signals. The LFP is the electrical potential recorded from deep brain tissues, such as the extracellular space in cortical regions. It is typically obtained by filtering the raw continuous data from 0.3 Hz to 500 Hz. Since LFPs mainly reflect the summation of electrical activities surrounding the recording site, they may be less sensitive to small movements or the loss of detectable neurons near the electrode tips [24]. Additionally, the extracellular medium, i.e., portions of cortical tissue between the electrode tip and neurons, acts as a low-pass filter. Consequently, the amplitude of high-frequency action potentials is naturally attenuated with distance. As a result, extracellular action potentials are only reliably detected for electrodes positioned near the recorded neuron. In contrast, low-frequency synaptic potentials attenuate less with distance, allowing them to propagate over larger distances in the extracellular medium. These low-frequency signals may even be recordable as far as the surface of the scalp, which is used in electro-encephalography (EEG) studies.

While early studies on LFPs primarily focused on the temporal analysis of a population of LFP signals, time-domain analyses of LFPs have demonstrated inferior performance for decoding neural information and control. One commonly-employed feature of interest is band-limited power, which can be computed over specific frequency bands, such as delta

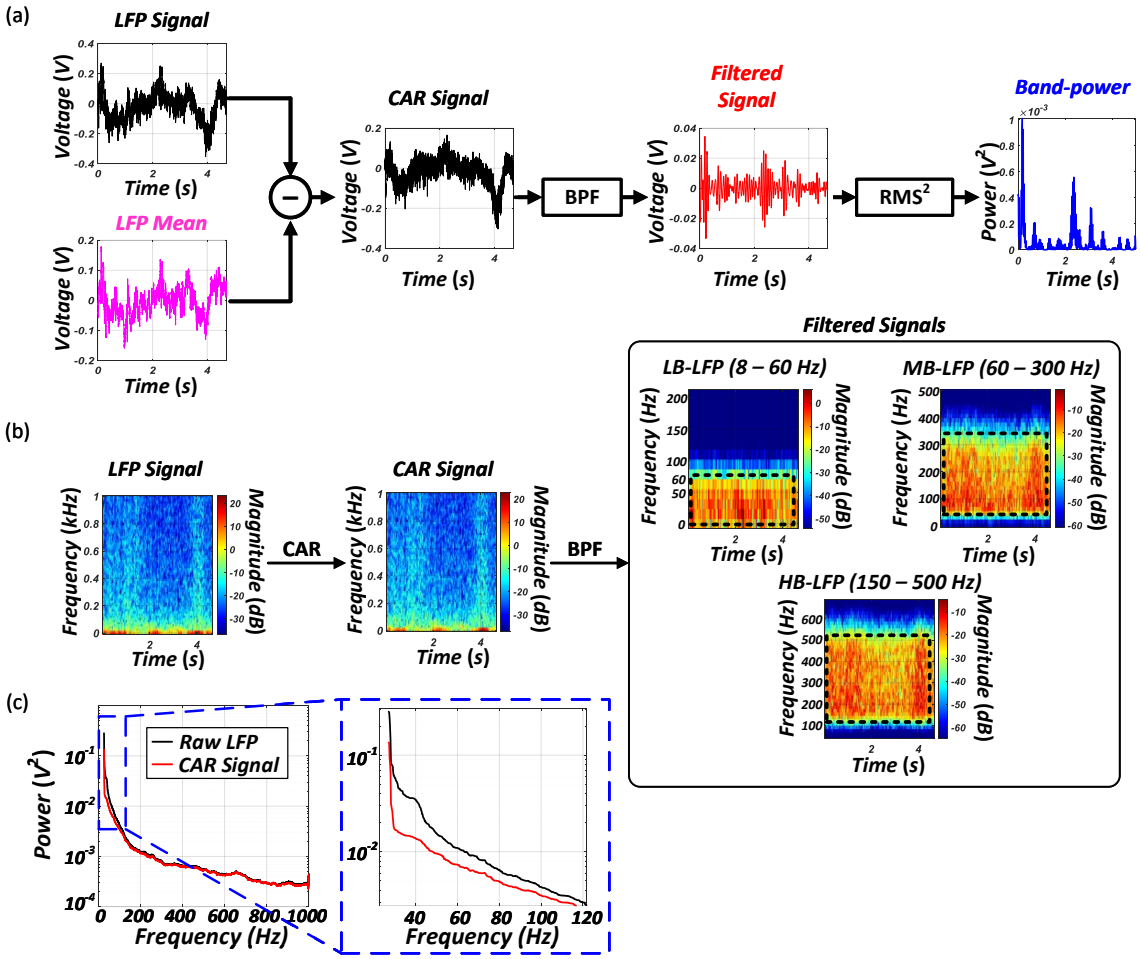

Fig. 2: (a) The block diagram of the in vivo LFP-based feature extraction pre-processing. (b) The time-frequency power spectrum of the neural signals during different phases of the feature extraction pre-processing. (c) The spectral power of the raw LFP and the CAR signal.

(1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz), beta (12–30 Hz), gamma (30–80 Hz), or high-frequency (80–500 Hz). It has been demonstrated that the spectral power in the high-frequency (HF) range of 60–200 Hz reflects the summation of spiking signals from the neuronal population near the recording electrode. This range may contain more sub-threshold information that is lost during the spike detection process. Since HF-LFP activity is correlated with neuronal firing rates (multi-unit spike count) [26], [27], it may provide particularly information-rich signals for predicting movement parameters, such as kinematic representations of prosthetic movement variables, even in the absence of clear spiking activity. For instance, a study in [13] reported that various movement intentions, such as the imagined end-point, trajectory, and type of movement, could be reliably predicted from the LFP signals. Additionally, LFPs are considered a more stable neural signal compared to the single-unit and multi-unit activities, which often suffer from electrode drift, neuron drop-out [28], and scarring from the electrode-tissue interface [25].

Another advantage of LFPs over action potentials is that LFPs can be sampled and processed at significantly lower rates (approximately 1 kHz) compared to spike events being digitized at sampling rates of 10 – 30 kHz for accurate spike detection [29]. A reduced sampling rate has a significant impact on the power consumption of implantable interfaces, potentially increasing the lifespan of implanted devices [29]. Additionally, there is no need for computationally-daunting spike detection and sorting algorithms, which consume 0.6 – 4 μW of power per channel [20], [30]. BCIs based on LFPs may also not require the use of deeply penetrating microelectrodes, significantly reducing the chance of tissue scarring and potentially increasing the longevity of the recorded neural signals [25]. Therefore, instead of relying on isolating individual action potentials within the 500–5000 Hz band, the low-frequency components ($<$ 500 Hz) of a recorded neural signal can be employed for a stable BCI signal processing [31].

III. Design of LFP-based RNN Brain Switches

To reduce the significant power consumption of the in vivo circuitry during non-BCI activities, a neurally-controlled brain-switch can be utilized. This switch enables or disables in vivo neural signal processing and wireless transmission of chosen features for in silico neural decoding. Brain-switch

algorithms operate on the assumption that the recorded neural signals exhibit various behaviors during different mental states. For instance, EEG-based brain-switches may detect an increase in the 1 Hz – 4 Hz band power [32], [33], while ECoG-based brain-switches may detect a change in the power of the specific frequency bands for motor execution/imagery [17]. For spike-based neural signal processing, the spike counts during a time interval [34], as well as changes in the firing rates [18], have been considered as underlying features for reported brain-switch algorithms. It has been previously reported that a small subset of recording channels exhibit satisfactory variations in the MUAs, allowing brain-switch algorithms to reliably detect transitions in mental state [20].

Due to the intracortical recording setting, LFPs offer a higher signal-to-noise ratio compared to ECoG and EEG. Moreover, the relaxed requirements of LFP signal acquisition and processing make them a viable modality for brain-switch algorithms, significantly reducing the power consumption compared to spike-based systems. To analyze the feasibility of using LFP features for brain-switch models, we employed two publicly-available neural datasets, datasets I (I140703) and II (L101210) [35]. These datasets contain raw recordings from the motor cortex (M1 Region) of two Rhesus macaque monkeys using an implanted 96-channel Utah Array. The recordings were initially sampled at 30 KHz and then downsampled to 10 KHz. Further downsampling to 2 kHz was performed to extract low-frequency LFPs. During the experiments, the monkeys performed a cued reach and grasp task to displace an object. The monkeys were presented with a series of cues indicating that the beginning of a trial and specifying one of four combinations of grip and displacement forces to use. Additional details of the animal experiments are described in [35].

A. LFP Feature Extraction Unit

After signal acquisition and digitization, the common average reference (CAR) is conventionally employed to effectively remove the common noise shared across electrodes. The CAR is given as $\hat{\mathbf{X}}[n] = \mathbf{X}[n] - \mu[n]$, where $\mathbf{X}[n]$ denotes the multi-channel neural signal, and $\mu[n]$ denotes the mean over all recording channels at time n. The CAR enhances the SNR and the interpretation of neural activity. Although the CAR is computed over all recording electrodes in most analyses, practical settings typically limit the CAR to a local subset of recording channels, as its impact on the system-level performance is negligible [36]. Following CAR, spectral power is computed for three frequency bands: the low-band LFP (LB-LFP) (8 – 60 Hz), mid-band LFP (MB-LFP) (60 – 300 Hz), and high-band LFP (HB-LFP) (150 – 500 Hz). Spectral power is commonly calculated using Fourier or Morlet wavelet transforms, providing time-frequency decomposition of signals.

A less computationally complex approach is to use a series of bandpass filters with cutoff frequencies for LB-LFP, MB-LFP, and HB-LFP. Once the signals are filtered, the spectral power within each frequency band of interest can be estimated using envelope detection schemes [18] or fast Fourier transform [37]. Applying bandpass filtering allows for attenuating signal components outside the desired frequency range, eliminating the need for Fourier transform to compute the spectral power over all frequencies. Fig. 2(a) illustrates the estimation of total power contained within a specified frequency range (band power) using the square of the root mean-square (RMS) of the bandpass filtered LFP signal in the time-domain [38]. Fig. 2(b) illustrates the time-frequency plots of the raw LFP signal, the output of the CAR, and the output of the bandpass filter (BPF) over the LB-LFP (8 – 60 Hz), MB-LFP (60 – 300 Hz), and the HB-LFP (150 – 500 Hz). The time-frequency plots in Fig. 2(b) demonstrate that the employed BPF effectively isolates the signal within the specified frequency range. Following the implementation of CAR, the multi-channel noise is reduced, leading to a decline in low-frequency power, as depicted in the magnified view presented in Fig. 2(c). For Channel 1 of Dataset I, depicted in Fig. 2, the average power of the raw LFP within 0 – 120 Hz is 16 mV^2, while after CAR, the average power in the same frequency band is reduced to 8.2 mV^2.

B. RNN Model Evaluation

While brain-switch software models may use all recording channels as the input to the algorithm, one opportunity for reducing energy dissipation is to employ a subset of recording channels. This, however, comes at the expense of additional pre-processing and a learning phase. This approach is viable for spike-based iBCIs, where specific neurons or an ensemble of neurons may behave differently during the intention onset. However, in the case of LFPs representing the activity of neuronal populations, it might be possible to utilize either a set of arbitrary recording channels or even a single recording channel as the input for the brain-switch algorithm.

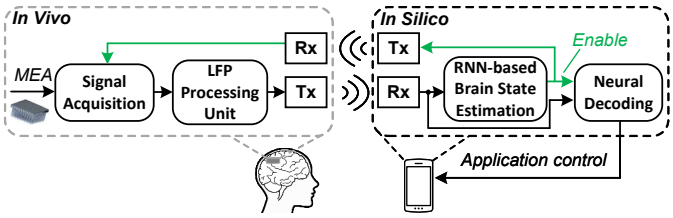

Fig. 3: The system-level block diagram of the designed brain-switch consisting of the in vivo LFP processing unit and in silico brain-state estimation.

In our previous work [20], we found that machine learning (ML) algorithms, specifically relatively small recurrent neural networks (RNNs) using neural spikes as input performed well for brain-state detection. In this study, we implement an LFP-based brain-switch RNN, as depicted in Fig. 3. The brain-switch consists of two units: an in vivo processing system for neural signal processing and LFP feature extraction, and an in silico module that receives wirelessly transmitted features and executes the RNN-based brain-switch model. When the BCI is in the "inactive" state, one recording channel is employed for in silico brain-state estimation. However, when the BCI is in the "active" state, all or a subset of recording channels can be employed for robust neural decoding. Therefore, the output of

the brain-switch is utilized to control both the in vivo signal acquisition as well as in silico neural decoding. It is important to note that the neural signal features used for decoding may differ from those used for the brain-state estimation. In vivo neural signals are first enhanced using the CAR. Then an 8-th order bandpass filter, along with a squared RMS module is used to estimate the power within each frequency band. where the number of spikes observed within this time period make up the multi-unit activities [29], [39]. Ranges up to 50 ms have been reported for iBCIs targeting the SBP [21]. The RMS is computed over 10 ms windows as it falls within the range of commonly employed iBCI configurations.

We evaluate the performance of four alternative RNN-based models with different recurrent cell architectures: standard RNN, long short-term memory (LSTM), gated recurrent unit (GRU), and quasi-recurrent neural nework (QRNN) [40]. Each RNN cell accepts the previous layer inputs and employs a self-recurrent connection to learn temporal dependencies within the data. The standard RNN cell transmits its hidden state directly. In contrast, the LSTM and GRU cells utilize internal non-linear functions to regulate the amount of data propagated over time. The QRNN employs a convolution operation over time instead of directly propagating its internal state, with its temporal receptive field width dependent on the convolutional kernel width. For performance comparison among the alternative RNN cells, we devised a RNN brain-switch that accepts the LFP bandpower of a single channel to predict the moment of intention. First, the single channel is passed to a fully-connected dense layer with 16 units, where its output connects to an RNN-based layer to learn temporal features within the signals. For a fair comparison among the different cell types, the RNN-based layers have about the same number of parameters. The model with the simple RNN cells employs 62 units, the LSTM employs 34 units, the GRU employs 32 units, and the QRNN employs 25 units. Finally, the outputs of the RNN-based layer passes through a fully-connected dense layer that performs the final regression and predicts the current intention state using a non-linear sigmoid function.

Each of the RNN-based layers has about 4800 parameters and is trained using the mean squared error loss metric and the Adam optimizer. Each is trained using 10-fold cross-validation to provide a fair estimate of its performance. The performance of the designed brain-switch RNN models is measured using the F-score metric given as $F = 2T_P/(2T_P + F_N + F_P)$, where T_P denotes the number of correctly detected intentions, F_N denotes the number of non-detected intentions, and F_P denotes the number of false positive intentions [41]. Figs. 4 (a) and (b) show the boxplots of the F-score of the four alternative models for the first recording channel of Dataset I and II, respectively. It can be seen that the GRU-based model provides the most consistent performance over both datasets and training iterations compared to the alternative models. In other words, while the LSTM and QRNN can perform well, whether a given training session will yield an acceptable performance is not deterministic. However, for a well performing training session, the training time is also of importance. Fig. 5(a) and (b) show the variation of F-score of the different models over the training period. In general, the GRU-based model achieves a relatively high performance considerably faster than the other models and is also shown to provide stable performance over time compared to the other recurrent neural network models. Therefore, we employ the GRU-based architecture, as shown in Fig. 6, to realize the brain-switch.

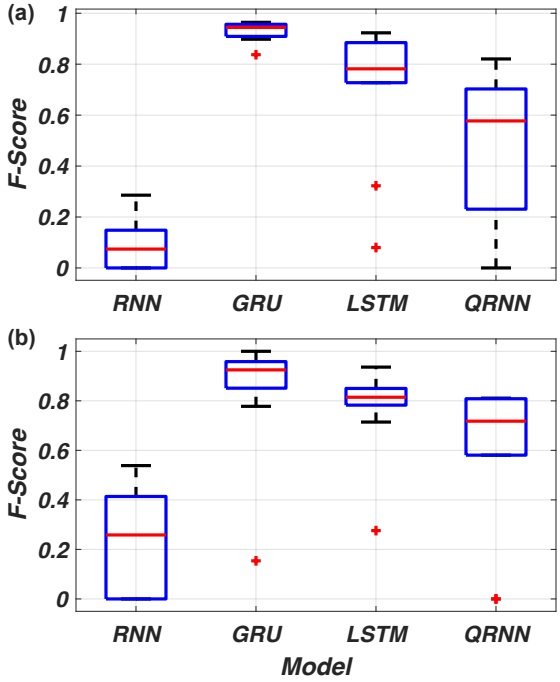

Fig. 4: The boxplots of the F-score of the four alternative RNN-based models over the first recording channel of (a) Dataset I and (b) Dataset II.

C. MEA Channel Selection

To evaluate the performance of the designed brain-switch scheme, the impact of channel and frequency band selections was analyzed. For each recording channel and frequency band, the data was normalized as $\hat{X_{fC}} = (X_{fC} - \mu_{fC})/\sigma_{fC}$, where μ_{fC} and σ_{fC} denote the mean and standard deviation of signal X of channel C filtered over frequency band f. Deep learning algorithms conventionally rely on a training and testing subset using 70% and 30% of the dataset, respectively. Compared to deep learning datasets that typically have millions of samples for training and testing, a common limitation in the field of computational neuroscience is that training data is relatively sparse. In our work, we have used only the first 30% of the dataset for training, which is consistent with our previously published work [10] to emulate this limitation. It is important to note that using less training data and more testing data implies that our model is not as intensely affected by overfitting. Dataset I and II consist of 153 and 149 trials, respectively. For each trial, the moment the monkey releases a switch is considered as the brain state transition from inactive to active, as used in our previous work [20]. However, to avoid detecting the intention during movement-related potentials when the animal releases the switch, a small randomized jitter

This article has been accepted for publication in IEEE Transactions on Biomedical Circuits and Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBCAS.2024.3396115

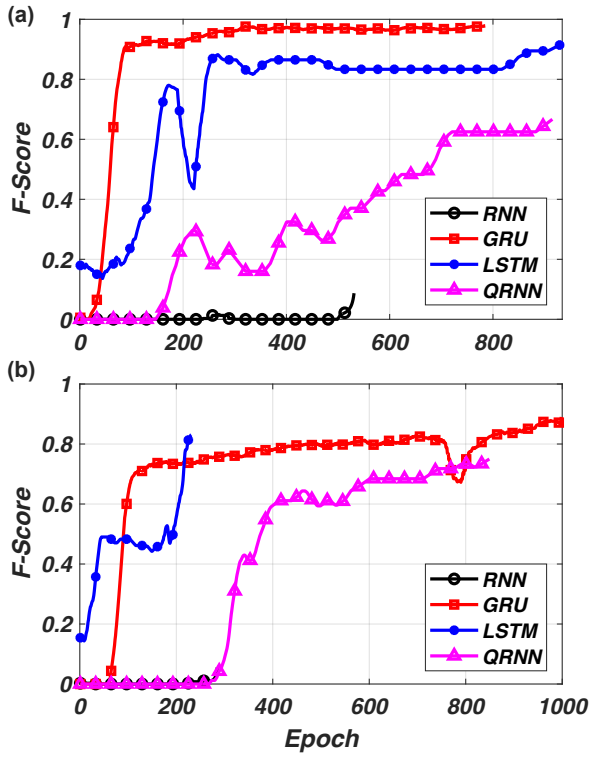

Fig. 5: The variation of F-score over the training epochs for (a) Dataset I and (b) Dataset II.

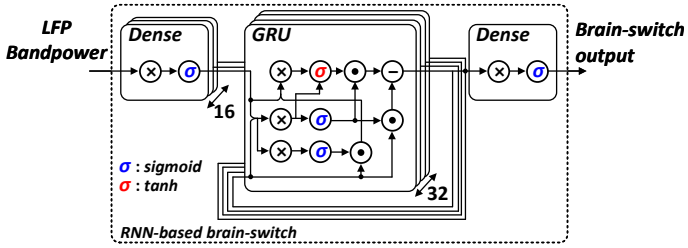

Fig. 6: The block diagram of the designed GRU-based LFP brain-switch.

time ($\mu = 14 \pm 8$ ms) was subtracted from the switch release time, ensuring that the state transition is detected during the movement planning phase of the trial. The value used for the jitter time has no particular significance other than ensuring that the RNN is trained to predict a brain-state transition prior to a switch released by the subject. The jitter time is modeled as a random variable so that the RNN would appropriately learn to perform the state transition based on the neural signal input patterns rather than performing a brain-state transition at the same time for any given trial. This approach also ensures that the model is only trained on neural signals that occur during the grasp planning stage of the subject's movement.

The designed RNN model was trained on all combinations of recording channels and frequency bands, and the F-Score probability distributions for Datasets I and II are shown in Figs. 7(a) and (b), respectively. It can be seen that the performance of the designed brain-switch model is agnostic to recording channel and frequency band over both datasets,

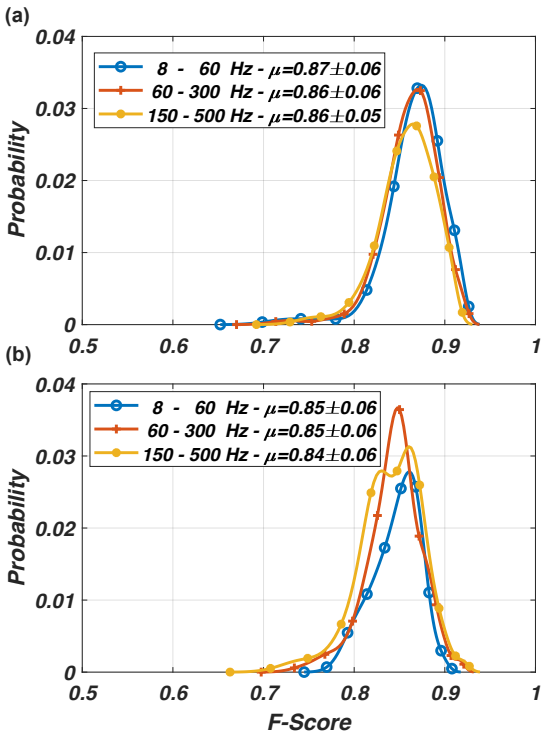

Fig. 7: The F-score probability distributions for (a) Dataset I and (b) Dataset II over the three frequency bands and all recording channels.

with a consistent F-Score performance of 0.86 ± 0.06 for Dataset I and approximately 0.80 ± 0.07 for Dataset II. This finding suggests that any particular channel may offer useful information for effective brain-state estimation. However, some channels may have lower performance than others. While beyond the scope of this work, one may employ a mean bandpower threshold or other relevant metrics for screening candidate channels. In our previous realization of the brain-switch, which utilized multi-unit activities in the form of spike threshold crossings, we required multiple channels along with a complex channel selection process for sufficiently accurate brain-state estimation [20]. However, we found that a single recording channel of LFPs, which embodies the activity of larger populations of neurons, provides relatively accurate brain-state estimation. This is potentially due to the fact that in the spike-based approach, removing low-frequency components to obtain the action potentials isolates the recorded neural signals to those of only neurons near the electrode tips, which may encode more specific information than that of an ensemble of neurons surrounding a larger region around the electrode.

D. CAR and Filter Considerations

The results depicted in Figs. 7(a) and (b) prompt the question of whether performing the CAR is necessary at all. CAR is commonly employed in LFP studies to remove noise that is common across the recording electrodes. In practice, removing common noise from multi-channel recordings is an effective approach for reducing the impact of redundant

This article has been accepted for publication in IEEE Transactions on Biomedical Circuits and Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBCAS.2024.3396115

features on the decoding algorithm. However, as discussed in Section II, our findings indicate that data from a single channel is adequate for the brain-switch operation. Therefore, the CAR may not be necessary, and the noise may be sufficiently removed by the bandpass filter. Additionally, in the context of efficient hardware realization, removing the CAR would significantly reduce the in vivo signal processing. Computing the CAR would require $2N-1$ additions and one multiplications per clock cycle, which can be expressed as $2N+1$ operations when assuming the complexity of multiplication as twice that of an addition [42]. Considering a clock rate of 2 kHz and a 96-channel Utah electrode array, the CAR would require 386 kOps/second.

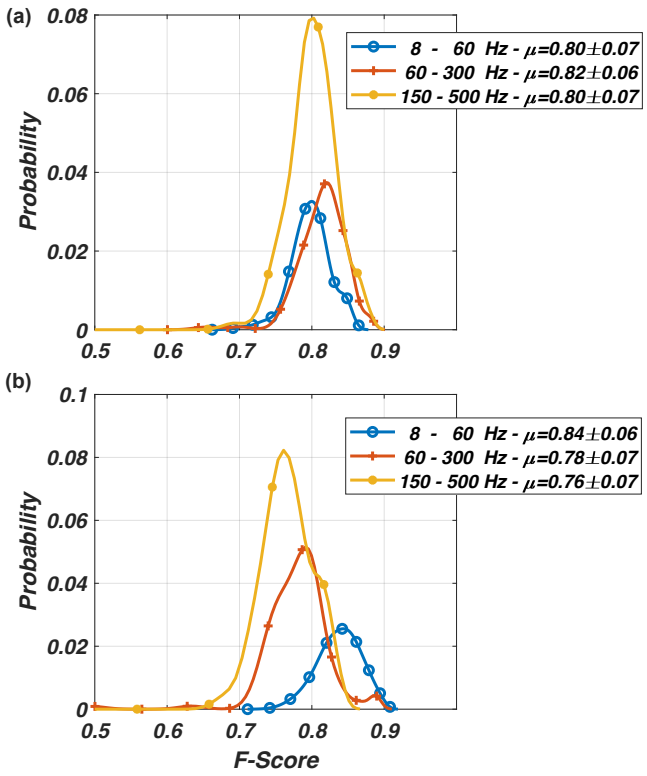

Fig. 8: The F-score probability distributions for (a) Dataset I and (b) Dataset II over the three frequency bands and all recording channels when removing CAR.

Figs. 8(a) and (b) show the F-Score probability distributions over Datasets I and II, respectively, and over all combinations of recording channels and frequency bands, without employing the CAR. It can be seen that for Dataset I, there is no significant degradation in F-score compared to those with the CAR for all three frequency bands. For Dataset II, the performance of each band varies, with the LB-LFP (8 – 15 Hz) outperforming the other two bands. Therefore, we employ the band-limited power of the LB-LFP as the input feature for the RNN-based brain-switch algorithm.

Previously published work mainly employ two-sided non-causal filters as they may better preserve the amplitude and shape of neural signals while not imposing phase distortions [43]. Causal filter realizations, however, are deemed to be preferred for in vivo realizations due to their reduced computational complexity and memory requirements. Fig. 9 shows the spectral power density of the first channel of Dataset I over the 8 – 15 Hz frequency band for both non-causal and causal filtering with various filter orders. It is evident that the lower-order filters do not limit the signal to the desired frequency band as accurately as the higher-order filters. While it may appear that the second-order causal filter spectrum shown in Fig. 9(d) may not be suitable, as it has a relatively poor frequency selectivity, we investigate its impact on the performance of the designed brain-switch algorithm by training the RNN model using both non-causal and causal filters of orders 2, 4, 6, and 8 and the LB-LFP bandpower for the first channel of Datasets I and II as input. Figs. 10(a) and (b) confirm that there is no significant performance variation when employing either non-causal or causal filters. Additionally, it is observed that for Dataset I, the lower-order filters outperform the higher-order filters, employing either non-causal or causal realizations. This suggests that the sharp roll-off characteristic of the higher-order filters is not a strict requirement for the designed RNN-based brain-switch. Therefore, we employ a second-order causal filter without CAR for in vivo realization. It should be noted that the performance of the proposed GRU-based brain state estimation is dependent on the dataset, as is the case for all machine learning-based algorithms. The performance of the proposed approach, and other similar kinematic decoding tasks, relies on the activation of relevant brain regions. Prior to surgical implantation of microelectrodes, the neural activity of the brain regions are observed using positron emission tomography (PET) or functional magnetic resonance imaging (fMRI). [44], [45].

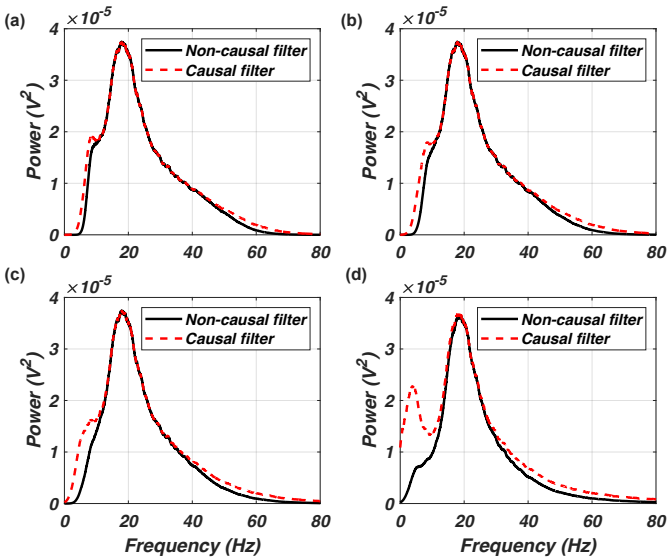

Fig. 9: The spectral power of the LB-LFP of non-causal and causal filters with filter order (a) 8, (b) 6, (c) 4, and (d) 2 over the first channel of Dataset I.

IV. Hardware Architecture of the Designed LFP-based Brain Switch

Fig. 11 shows the block diagram of the in vivo signal processing for the LFP-based feature extraction, i.e., LFP

This article has been accepted for publication in IEEE Transactions on Biomedical Circuits and Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TBCAS.2024.3396115

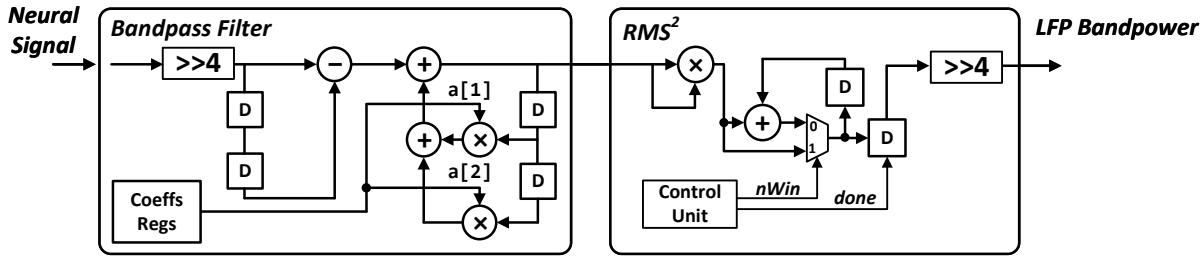

Fig. 11: The block diagram of the in vivo signal processing for estimating LFP bandpower.

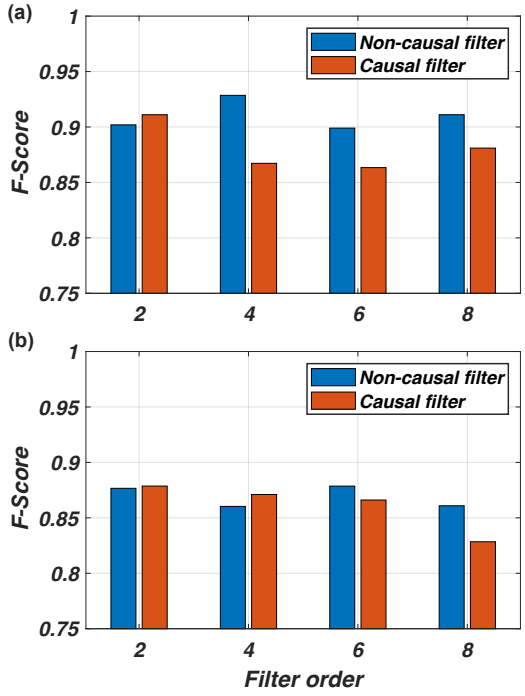

Fig. 10: The F-score of the LFP-based RNN brain-switch model using various filter orders and filter causalities for the LB-LFP and the first channel of (a) Dataset I and (b) Dataset II.

bandpower estimation. A single channel of the MEA is band-pass filtered at a rate of 2 kS/s. The bandpower is estimated by computing the RMS squared over non-overlapping windows of 8 ms (i.e., 16 samples at 2 kS/s). The bandpass filter is designed using a second-order infinite-impulse response (IIR) Butterworth filter structure [46]. The passband filter coefficients for $8 - 60$ Hz were computed using the Filter Design Toolbox in Matlab as $b_0 = 1$, $b_1 = 0$, $b_2 = -1$, $a[0] = 1$, $a[1] = -1.84$, and $a[2] = 0.848$. The filter input scaling factor of 0.075 is approximated as an arithmetic right shift by four bit positions, equivalent to multiplying by 0.0625. To find an optimal tradeoff between area utilization and power consumption, in addition to the direct biquad structure, the time-multiplexed IIR filter structures with the folding factors of 2 and 3, as shown in Figs. 12(a) – (b), respectively, were designed and implemented. The direct structure is required to operate at the minimum rate of 2 kHz while the designs with the folding factors of 2 and 3 are required to operate at 4 kHz and 6 kHz, respectively. A control unit, not shown, is used to generate appropriate control signals.

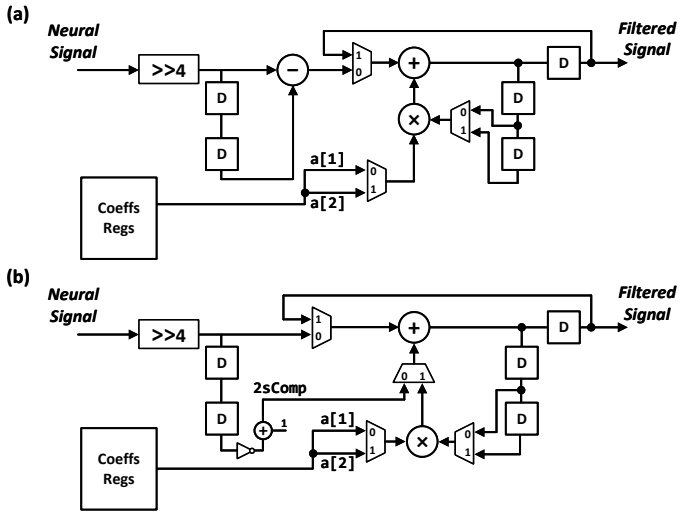

Fig. 12: The block diagram of the designed band-pass filter with the folding factors of (a) two and (b) three.

The filter output is passed to the *RMS*^{2} unit, as shown in Fig. 11. The *control unit* asserts the signal *nWin* every 8 ms, setting the accumulation register to the squared input sample. On the 16-th input from the filter (i.e., after 8 ms), the *control unit* asserts the signal *done* to latch the output of the accumulator into the output register, which is arithmetically right shifted by four bits to compute the mean over the window. The squared RMS is used as the estimate of the bandpower where no square root operation is required.

The computations of the filter and *RMS*^{2} unit are performed in the fixed-point format Q(WI.WF), where WI and WF denote the integer and fractional wordlengths of the signals. Software simulations of the in vivo modules verified that the filter's internal accumulator requires 17 bits in Q(5.12) format and the filter output requires 16 bits in Q(3.13) format. Since the *RMS*^{2} performs a square operation and a right shift for computing the mean, the full-precision computation would require at least 30 fractional bits. To find the optimal number of fractional bits necessary to achieve performance (F-Score) comparable to that of the software simulations shown in Figs. 10, various bit-true simulations were performed. Fig. 13 shows the F-Scores of the designed RNN-based brain-switch algorithm over different number of fractional bits for representing the RMS bandpower.

The analysis over the two datasets showed that a minimum of eight fractional bits offers sufficient accuracy.

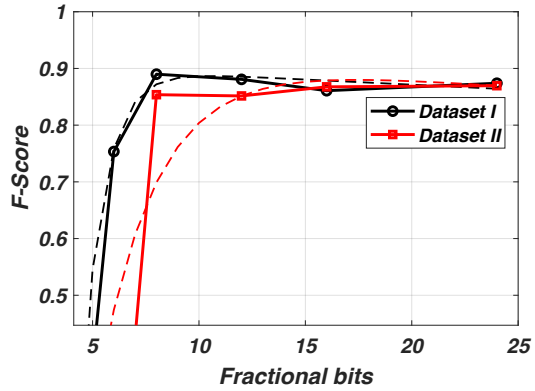

Fig. 13: The variation of the F-Score of the designed RNN brain-switch over various number of fractional bits used for representing the bandpower, with the dashed lines indicating the trend.

The logic synthesis of the in vivo design shown in Fig. 11 was implemented using Synopsys Design Compiler, and the place and route was performed using Cadence Innovus in a standard 180-nm CMOS process. The design was also synthesized for both folding factors of the bandpass filter datapath. The power consumption was estimated by employing both datasets, along with the toggling rate information of all signals stored in a Switching Activity Interchange Format (SAIF) file. Table I gives the area utilization and power consumption of the designed LFP bandpower estimator using the direct structure and using two folding factors. It is shown that the smallest area and the least power consumption are achieved with the direct-form IIR filter structure. While the folding factors of two and three for the bandpass filter datapath have fewer adders and multipliers, the multiplexers require sign and zero extension of inputs with fewer bits to match those of the accumulation registers, which in turn require more area. While the LB-LFP offered the best performance for both employed datasets, it may be beneficial to support MB-LFP and HB-LFP bands as well for practical applications (e.g., neural decoding). The design was therefore modified to support all three sub-bands. The power consumption of the LFP-based feature extraction unit with the programmable frequency band for different folding factors is given in Table I. Similar to that of the ASIC for the LB-LFP, the direct structure (i.e., the folding factor of one) consumes the least power among the three realizations for a marginally larger silicon area.

The ASIC layouts of the synthesized in vivo LFP processing unit, including the IIR filter and the RMS bandpower unit, for (a) the LB-LFP and (b) the programmable frequency band in a standard 180-nm CMOS process are shown in Fig. 14. The layout in Fig. 14(a) computes the bandpower of the LB-LFP and occupies 0.089 mm^2 of silicon area and consumes 91.87 nW of power from a 1.8 V supply while operating at 2 kHz. The layout in Fig. 14(b) computes the bandpower of any one of the LB-LFP, MB-LFP, and HB-LFP bands based on a programmable band-selection input, and occupies 0.117 mm^2 of silicon area and consumes 109.9 nW of power from a 1.8 V supply, while operating at 2 kHz. The power consumption was estimated by simulating the netlist after placement and routing and accounting for the switching activity of all signals.

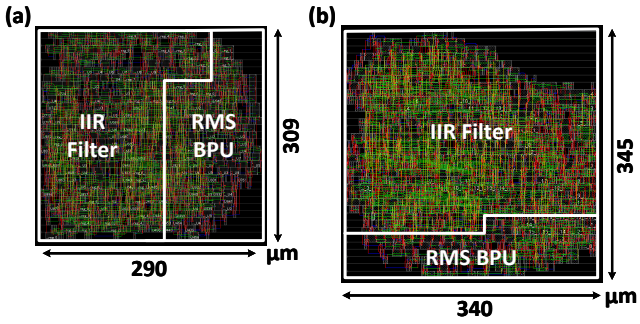

Fig. 14: The ASIC layouts of the synthesized in vivo LFP-based feature extraction units for (a) the LB-LFP and (b) the programmable frequency band in a standard 180-nm CMOS process.

Table II gives the characteristics and implementation results of various in vivo feature extraction ASICs. Due to the relatively high temporal resolution of action potentials, analog front-ends (AFEs) for spike-based BCIs must operate at a significantly higher sampling rate and bandwidth than those employing LFPs. First, the power consumption of the low-noise amplifier (LNA) is proportional to the width of the frequency band of interest [29], which for LFPs is about five times smaller than that for spikes. Secondly, conventional AFEs employ successive approximation register (SAR) analog-to-digital converters (ADCs), whose power consumption is relatively linearly proportional to the sampling rate [50], [51], which again are far more efficient for LFPs. Thus, for a fair comparison with the previously published works, the comparison Table II only accounts for the digital backend of all works. In our previous work [20], we designed and implemented an RNN-based brain-switch algorithm employing MUA features. Compared to the designed LFP-based brain switch, the MUA-based design requires a correlation-based learning phase to find the optimal recording channels for detecting the brain's state transitions. Employing the same datasets, the MUA-based brain-switch algorithm achieves a marginally improved F-score of 0.89 compared to the LFP-based brain-switch F-score of 0.84, with a power consumption approximately five times higher. In [47], the authors present an analog implementation of a spike detection unit that performs energy-based thresholding to generate MUAs. In [48], the authors target the low-frequency SBP and achieve a low power consumption of 212 nW per channel. Also targeting the SBP, the design in [49] reports a power consumption of 3.6 μW per recording channel for obtaining the SBP. The proposed LFP-based ASICs, as given in Table II, consume significantly lower power for extracting the LFP features. While it remains to explore how well the SBP can be employed for deteciting the brain-state transitions, due to the LFP's lower frequency filtering and acquisition requirements, the LFP is

TABLE II: The characteristics and implementation results of various feature extraction ASICs.

Work	Ours (LB-LFP)	Ours (Prog.)	Ours[20]	[47]	[48]	[49]
Signal modality	LFPs		MUAs	MUAs	SBP	SBP
Technology (nm)	180	180	180	180	180	180
Supply voltage (V)	1.8	1.8	1.8	1.8	1.55	0.625
Area per channel (mm^2)	0.09	0.12	0.03	0.03	0.07	–
Power per channel (nW)†	91.8	109.9	630	1500	212	3660

† Normalized to a 180-nm CMOS process with a 1.8 V supply and accounting only for digital backend.

TABLE I: The area utilization and power consumption of the designed and implemented LFP-based feature extraction unit.

Folding Factor	**Area (mm^2)**	**Power (nW)**
1	0.09	91.8
2	0.10	192.5
3	0.10	160.6
Programmable Frequency Band		
1	0.12	109.9
2	0.12	247.2
3	0.11	279.6

well positioned as an effective signal modality for brain-state monitoring and behavioral decoding.

One important consideration is the total power consumption of the proposed system, accounting for the signal acquisition analog front-end (AFE) circuitry, digital signal processing, and wireless transmission overhead. Conventional AFE circuitry, such as low-noise amplifiers (LNAs) and analog-to-digital converters (ADC), requires an average of 7.4 μW of power per channel for spike-based systems employing sampling rates on the order of 10 kS/s [52]. As both the bandwidth and sampling rate of the LFPs and SBPs is roughly one-fifth of that for spike-based BCIs, the signal conditioning AFE for LFPs and SBPs would consume approximately 1.5 μW per recording channel.

The wireless transmission power depends largely on the output bitrate of the in vivo circuitry, as well as the temporal resolution of the signal features. For spike-based systems, a commonly employed range of temporal resolutions is within the range of 1 ms to 10 ms, requiring an output data rate between 0.4 - 1.0 kbps, i.e., one bit per millisecond or $\lceil \log_2 10\ \mathrm{ms} \rceil = 4$ bits per millisecond assuming a one millisecond spike refractory period. The temporal resolution of the proposed LFP feature extraction unit is 10 milliseconds, with 9 bits (Q(1.8)) per estimated bandpower sample and requires 0.9 kbps. Spike-based systems consume between 0.6 and 1.78 μW of power for obtaining MUAs. Given a wireless transmission power of 158 pJ/bit [53], spike-based systems require between 8.06 μW and 9.26 μW per recording channel, accounting for the entire signal acquisition and processing chain, and wireless transmission. The designed LFP-based feature extraction system would require a total of 1.75 μW per recording channel. Based on our previous work implementing on a spike-based brain-switch targeting MUAs, we found that a minimum of 8 recording channels were required for sufficient brain-state transition detection, which would require a total of 64.48 μW of power. In contrast, the designed LFP-based brain-switch requires only a single channel and consumes over 97% less power with only 1.75 μW.

Considering that a single channel of the LFP is sufficient for brain state estimation, an alternative system configuration is to bypass the LFP feature extraction in vivo and transmit the LFP signal directly. Assuming a signal resolution of 16 bits and a sampling rate of 2 kS/s, the required output data rate for a system transmitting the LFP signal directly is 32 kbps. Given the transmission energy of 158 pJ/bit, this would require 5.05 μW of power, over 2.8$\times$ more power than that of the proposed system. For transmitting the LFP signal directly, the increased data rate becomes the primary factor causing increased wireless transmission power requirements. To achieve a comparable power consumption to the proposed design, the LFP signal must be represented with fewer than 6 bits to consume less than 1.8 μW.

V. Conclusion

This article presented a brain-switch to estiamte when the user involves in the underlying BCI activity based on local field potentials (LFPs) and gated recurrent unit (GRU)-based neural networks. The designed switch is particularly imnportant in the practical realization of asynchronous brain-computer interfaces and offers a comparable or greater effectiveness than that of spike-based systems. It was shown that the performance of the machine learning (ML)-based algorithms for implementing LFP-based brain-switches may not be sensitive to specific recording channels as well as specific frequency bands. Compared to the previously reported brain-switch schemes in which exhaustive learning phases for optimal recording channels are performed prior to neural network training, the brain-switch performance is similar across multiple recording channels, significantly reducing the pre-processing computations. The synthesized LFP-based feature extraction unit in a standard 180-nm CMOS process was estimated to occupy 0.09 mm^2 of silicon area and consumes only 91.8 nW of power while operating at 2 kHz. Compared to the previously published works on in vivo neural signal feature extraction, the proposed LFP-based design consumes the least power and achieves a performance (F-Score) comparable to that of spike-based systems.

Data Availability

The data presented in this study are publicly available on G-Node at https://doi.org/10.12751/g-node.f83565, reference [35].

Acknowledgment

This work was supported by the National Science Foundation Award #2007131. We would like to thank Brochier et al. [35] for providing open access to Datasets I and II.

REFERENCES

[1] W. Wu, M. J. Black, Y. Gao, E. Bienenstock, M. Serruya, A. Shaikhouni, and J. P. Donoghue, "Neural decoding of cursor motion using a kalman filter," *Advances in neural information processing systems*, pp. 133–140, 2003.

[2] M. Serruya, N. Hatsopoulos, M. Fellows, L. Paninski, and J. Donoghue, "Robustness of neuroprosthetic decoding algorithms," *Biological Cybernetics*, vol. 88, no. 3, pp. 219–228, 2003.

[3] C. Gold, D. A. Henze, C. Koch, and G. Buzsaki, "On the origin of the extracellular action potential waveform: A modeling study," *Journal of Neurophysiology*, vol. 95, no. 5, pp. 3113–3128, 2006.

[4] M. S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials," *Network: Computation in Neural Systems*, vol. 9, no. 4, p. R53, 1998.

[5] S. Todorova, P. Sadtler, A. Batista, S. Chase, and V. Ventura, "To sort or not to sort: the impact of spike-sorting on neural decoding performance," *Journal of Neural Engineering*, vol. 11, no. 5, p. 056005, 2014.

[6] E. Stark and M. Abeles, "Predicting movement from multiunit activity," *Journal of Neuroscience*, vol. 27, no. 31, pp. 8387–8394, 2007.

[7] J. D. Wander and R. P. Rao, "Brain–computer interfaces: a powerful tool for scientific inquiry," *Current Opinion in Neurobiology*, vol. 25, pp. 70–75, 2014.

[8] E. R. Oby, S. Perel, P. T. Sadtler, D. A. Ruff, J. L. Mischel, D. F. Montez, M. R. Cohen, A. P. Batista, and S. M. Chase, "Extracellular voltage threshold settings can be tuned for optimal encoding of movement and stimulus parameters," *Journal of Neural Engineering*, vol. 13, no. 3, p. 036009, 2016.

[9] F. R. Willett, D. T. Avansino, L. R. Hochberg, J. M. Henderson, and K. V. Shenoy, "High-performance brain-to-text communication via handwriting," *Nature*, vol. 593, no. 7858, pp. 249–254, 2021.

[10] D. Valencia, J. Thies, and A. Alimohammad, "Frameworks for efficient brain-computer interfacing," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 6, pp. 1714–1722, 2019.

[11] M. Filippini, A. Morris, R. Breveglieri, K. Hadjidimitrakis, and P. Fattori, "Decoding of standard and non-standard visuomotor associations from parietal cortex," *Journal of neural engineering*, vol. 17, no. 4, p. 046027, 2020.

[12] S. Musallam, B. Corneil, B. Greger, H. Scherberger, and R. A. Andersen, "Cognitive control signals for neural prosthetics," *Science*, vol. 305, no. 5681, pp. 258–262, 2004.

[13] T. Aflalo, S. Kellis, C. Klaes, B. Lee, Y. Shi, K. Pejsa, K. Shanfield, S. Hayes-Jackson, M. Aisen, C. Heck *et al.*, "Decoding motor imagery from the posterior parietal cortex of a tetraplegic human," *Science*, vol. 348, no. 6237, pp. 906–910, 2015.

[14] E. Christensen, A. Abosch, J. A. Thompson, and J. Zylberberg, "Inferring sleep stage from local field potentials recorded in the subthalamic nucleus of parkinson's patients," *Journal of Sleep Research*, vol. 28, no. 4, p. e12806, 2019.

[15] A. R. Adamantidis, C. Gutierrez Herrera, and T. C. Gent, "Oscillating circuitries in the sleeping brain," *Nature Reviews Neuroscience*, vol. 20, no. 12, pp. 746–762, 2019.

[16] B. E. Kilavik, M. Zaepffel, A. Brovelli, W. A. MacKay, and A. Riehle, "The ups and downs of beta oscillations in sensorimotor cortex," *Experimental neurology*, vol. 245, pp. 15–26, 2013.

[17] J. J. Williams, A. G. Rouse, S. Thongpang, J. C. Williams, and D. W. Moran, "Differentiating closed-loop cortical intention from rest: building an asynchronous electrocorticographic bci," *Journal of Neural Engineering*, vol. 10, no. 4, p. 046001, 2013.

[18] J. J. Williams, R. N. Tien, Y. Inoue, and A. B. Schwartz, "Idle state classification using spiking activity and local field potentials in a brain computer interface," in *Engineering in Medicine and Biology Society*. IEEE, 2016, pp. 1572–1575.

[19] C.-H. Han, K.-R. Müller, and H.-J. Hwang, "Brain-switches for asynchronous brain–computer interfaces: A systematic review," *Electronics*, vol. 9, no. 3, p. 422, 2020.

[20] D. Valencia, G. Leone, N. Keller, P. P. Mercier, and A. Alimohammad, "Power-efficient in vivo brain-machine interfaces via brain-state estimation," *Journal of Neural Engineering*, vol. 20, no. 1, p. 016032, 2023.

[21] S. R. Nason, A. K. Vaskov, M. S. Willsey, E. J. Welle, H. An, P. P. Vu, A. J. Bullard, C. S. Nu, J. C. Kao, K. V. Shenoy *et al.*, "A low-power band of neuronal spiking activity dominated by local single units improves the performance of brain–machine interfaces," *Nature Biomedical Engineering*, vol. 4, no. 10, pp. 973–983, 2020.

[22] S. Katzner, I. Nauhaus, A. Benucci, V. Bonin, D. L. Ringach, and M. Carandini, "Local origin of field potentials in visual cortex," *Neuron*, vol. 61, no. 1, pp. 35–41, 2009.

[23] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents – EEG, ECoG, LFP and spikes," *Nature Reviews Neuroscience*, vol. 13, no. 6, pp. 407–420, 2012.

[24] S. Saha, K. A. Mamun, K. Ahmed, R. Mostafa, G. R. Naik, S. Darvishi, A. H. Khandoker, and M. Baumert, "Progress in brain computer interface: Challenges and opportunities," *Frontiers in Systems Neuroscience*, vol. 15, p. 578875, 2021.

[25] J. W. Salatino, K. A. Ludwig, T. D. Kozai, and E. K. Purcell, "Glial responses to implanted electrodes in the brain," *Nature biomedical engineering*, vol. 1, no. 11, pp. 862–877, 2017.

[26] D. A. Heldman, W. Wang, S. S. Chan, and D. W. Moran, "Local field potential spectral tuning in motor cortex during reaching," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 180–183, 2006.

[27] J. A. Perge, S. Zhang, W. Q. Malik, M. L. Homer, S. Cash, G. Friehs, E. N. Eskandar, J. P. Donoghue, and L. R. Hochberg, "Reliability of directional information in unsorted spikes and local field potentials recorded in human motor cortex," *Journal of Neural Engineering*, vol. 11, no. 4, p. 046007, 2014.

[28] J. E. Downey, N. Schwed, S. M. Chase, A. B. Schwartz, and J. L. Collinger, "Intracortical recording stability in human brain–computer interface users," *Journal of neural engineering*, vol. 15, no. 4, p. 046016, 2018.

[29] N. Even-Chen, D. G. Muratore, S. D. Stavisky, L. R. Hochberg, J. M. Henderson, B. Murmann, and K. V. Shenoy, "Power-saving design opportunities for wireless intracortical brain–computer interfaces," *Nature Biomedical Engineering*, vol. 4, no. 10, pp. 984–996, 2020.

[30] Y. Chen, B. Tacca, Y. Chen, D. Biswas, G. Gielen, F. Catthoor, M. Verhelst, and C. M. Lopez, "A 384-channel online-spike-sorting ic using unsupervised geo-osort clustering and achieving 0.0013 mm^2/ch and 1.78μw/ch," in *International Solid-State Circuits Conference*. IEEE, 2023, pp. 486–488.

[31] T. Milekovic, A. A. Sarma, D. Bacher, J. D. Simeral, J. Saab, C. Pandarinath, B. L. Sorice, C. Blabe, E. M. Oakley, K. R. Tringale *et al.*, "Stable long-term bci-enabled communication in als and locked-in syndrome using lfp signals," *Journal of Neurophysiology*, vol. 120, no. 7, pp. 343–360, 2018.

[32] Z. Bozorgzadeh, G. E. Birch, and S. G. Mason, "The LF-ASD brain computer interface: on-line identification of imagined finger flexions in the spontaneous EEG of able-bodied subjects," in *International Conference on Acoustics, Speech, and Signal Processing*, vol. 4. IEEE, 2000, pp. 2385–2388.

[33] J. F. Borisoff, S. G. Mason, and G. E. Birch, "Brain interface research for asynchronous control applications," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 160–164, 2006.

[34] N. Achtman, A. Afshar, G. Santhanam, M. Y. Byron, S. I. Ryu, and K. V. Shenoy, "Free-paced high-performance brain–computer interfaces," *Journal of Neural Engineering*, vol. 4, no. 3, p. 336, 2007.

[35] T. Brochier, L. Zehl, Y. Hao, M. Duret, J. Sprenger, M. Denker, S. Grün, and A. Riehle, "Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task," *Scientific Data*, vol. 5, no. 1, pp. 1–23, 2018.

[36] N. Ahmadi, T. G. Constandinou, and C.-S. Bouganis, "Impact of referencing scheme on decoding performance of lfp-based brain-machine interface," *Journal of Neural Engineering*, vol. 18, no. 1, p. 016028, 2021.

[37] F. J. Harris, *Multirate signal processing for communication systems*. CRC Press, 2022.

[38] J. Tamarkin, "A new proof of parseval's identity for trigonometric functions," *Annals of Mathematics*, pp. 541–547, 1926.

[39] M. M. Shanechi, A. L. Orsborn, H. G. Moorman, S. Gowda, S. Dangi, and J. M. Carmena, "Rapid control and feedback rates enhance neuroprosthetic control," *Nature communications*, vol. 8, no. 1, p. 13825, 2017.

[40] J. Bradbury, S. Merity, C. Xiong, and R. Socher, "Quasi-Recurrent Neural Networks," *International Conference on Learning Representations*, 2017.

[41] C. J. Van Rijsbergen, "Information retrieval," 1979.

[42] H. Mora-Mora, J. Mora-Pascual, J. M. García-Chamizo, and A. Jimeno-Morenilla, "Real-time arithmetic unit," *Real-Time Systems*, vol. 34, pp. 53–79, 2006.

[43] D. Valencia, P. P. Mercier, and A. Alimohammad, "In vivo neural spike detection with adaptive noise estimation," *Journal of Neural Engineering*, vol. 19, no. 4, p. 046018, 2022.

[44] H. Van Mier, L. Tempel, J. Perlmutter, M. Raichle, and S. Petersen, "Changes in brain activity during motor learning measured with pet: effects of hand of performance and practice," *Journal of Neurophysiology*, vol. 80, no. 4, pp. 2177–2199, 1998.

[45] D. M. Mehler, A. N. Williams, F. Krause, M. Lührs, R. G. Wise, D. L. Turner, D. E. Linden, and J. R. Whittaker, "The bold response in primary motor cortex and supplementary motor area during kinesthetic motor imagery based graded fmri neurofeedback," *Neuroimage*, vol. 184, pp. 36–44, 2019.

[46] V. Shinde, G. Jai Kumar, D. Valencia, and A. Alimohammad, "High-throughput and compact reconfigurable architectures for recursive filters," *IET Communications*, vol. 12, no. 13, pp. 1616–1623, 2018.

[47] E. Koutsos, S. E. Paraskevopoulou, and T. G. Constandinou, "A 1.5 μw NEO-based spike detector with adaptive-threshold for calibration-free multichannel neural interfaces," in *IEEE International Symposium on Circuits and Systems*, 2013, pp. 1922–1925.

[48] G. Atzeni, J. Lim, J. Liao, A. Novello, J. Lee, E. Moon, M. Barrow, J. Letner, J. Costello, S. R. Nason *et al.*, "A 260$\times$ 274 μm 2 572 nw neural recording micromote using near-infrared power transfer and an rf data uplink," in *Symposium on VLSI Technology and Circuits*. IEEE, 2022, pp. 64–65.

[49] H. An, S. R. Nason-Tomaszewski, J. Lim, K. Kwon, M. S. Willsey, P. G. Patil, H.-S. Kim, D. Sylvester, C. A. Chestek, and D. Blaauw, "A power-efficient brain-machine interface system with a sub-mw feature extraction and decoding asic demonstrated in nonhuman primates," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 16, no. 3, pp. 395–408, 2022.

[50] M. Yip and A. P. Chandrakasan, "A resolution-reconfigurable 5-to-10-bit 0.4-to-1 v power scalable SAR ADC for sensor applications," *IEEE Journal of Solid-State Circuits*, vol. 48, no. 6, pp. 1453–1464, 2013.

[51] H. Wang, X. Wang, A. Barfidokht, J. Park, J. Wang, and P. P. Mercier, "A battery-powered wireless ion sensing system consuming 5.5 nw of average power," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 7, pp. 2043–2053, 2018.

[52] J. Li, X. Liu, W. Mao, T. Chen, and H. Yu, "Advances in neural recording and stimulation integrated circuits," *Frontiers in Neuroscience*, vol. 15, p. 663204, 2021.

[53] J. Rosenthal and M. S. Reynolds, "A 158 pj/bit 1.0 mbps bluetooth low energy (BLE) compatible backscatter communication system for wireless sensing," in *Conference on Wireless Sensors and Sensor Networks (WiSNet)*. IEEE, 2019, pp. 1–3.

Patrick P. Mercier (S'04–M'12–SM'17) received the B.Sc. degree in electrical and computer engineering from the University of Alberta, Edmonton, AB, Canada, in 2006, and the S.M. and Ph.D. degrees in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, MA, USA, in 2008 and 2012, respectively.

He is currently a Professor in Electrical and Computer Engineering at the University of California San Diego (UCSD), where he is also the co-Director of the Center for Wearable Sensors and the Site Director of the Power Management Integration Center. His research interests include the design of energy-efficient microsystems, focusing on the design of RF circuits, power converters, and sensor interfaces for miniaturized systems and biomedical applications.

Prof. Mercier has published over 200 peer-reviewed papers, including 26 ISSCC papers, 34 JSSC papers, and several papers in high-impact journals such as Science, Nature Biotechnology, Nature Biomedical Engineering, Nature Electronics, Nature Communications, Advanced Science, and more. He has received numerous awards, including a Natural Sciences and Engineering Council of Canada (NSERC) Julie Payette fellowship in 2006, NSERC Postgraduate Scholarships in 2007 and 2009, an Intel Ph.D. Fellowship in 2009, the 2009 IEEE International Solid-State Circuits Conference (ISSCC) Jack Kilby Award for Outstanding Student Paper at ISSCC 2010, a Graduate Teaching Award in Electrical and Computer Engineering at UCSD in 2013, the Hellman Fellowship Award in 2014, the Beckman Young Investigator Award in 2015, the DARPA Young Faculty Award in 2015, the UC San Diego Academic Senate Distinguished Teaching Award in 2016, the Biocom Catalyst Award in 2017, the NSF CAREER Award in 2018, a National Academy of Engineering Frontiers of Engineering Lecture in 2019, the San Diego County Engineering Council Outstanding Engineer Award in 2020, the ISSCC Author Recognition Award in 2023, and the ECE Teacher of the Year award in 2023. He has served as an Associate Editor of the IEEE Transactions on Very Large Scale Integration (TVLSI), the IEEE Transactions on Biomedical Circuits and Systems (TBioCAS), and the IEEE Solid-State Circuits Letters. He is currently a member of the Executive Committee of ISSCC, and has served on the technical program committees for ISSCC, CICC, and the VLSI Symposium. Prof. Mercier was the co-editor of Ultra-Low-Power Short Range Radios (Springer, 2015) Power Management Integrated Circuits (CRC Press, 2016), and High-Density Electrocortical Neural Interfaces (Academic Press, 2019).

Daniel Valencia is a graduate researcher working in the VLSI Design and Test Laboratory in the Department of Electrical and Computer Engineering at the San Diego State University. He is currently a Ph.D. candidate in the joint doctoral program at San Diego State University and University of California, San Diego. His research interests include field-programmable gate arrays, brain-computer interfacing, and VLSI architectures for neural signal processing.

Amir Alimohammad is a Professor in the Department of Electrical and Computer Engineering at San Diego State University. He obtained his Ph.D. from the University of Alberta in Canada. Currently, he serves as an Associate Editor for Frontiers in Neuroscience and Neuromorphic Engineering. Moreover, he is a Member of the National Science Foundation (NSF) Center for Neurotechnology and a Senior Member of IEEE. His primary research focus revolves around digital VLSI design for brain-computer interfacing and wireless communication technologies.