Predicting Single-Unit Activity from Local Field Potentials with LSTMs

Oscar W. Savolainen, *Student Member, IEEE*, Timothy G. Constandinou, *Senior Member, IEEE*

***Abstract*— This paper investigates to what extent Long Short-Term Memory (LSTM) decoders can use Local Field Potentials (LFPs) to predict Single-Unit Activity (SUA) in Macaque Primary Motor cortex. The motivation is to determine to what degree the LFP signal can be used as a proxy for SUA, for both neuroscience and Brain-Computer Interface (BCI) applications. Firstly, the results suggest that the prediction quality varies significantly by implant location or animal. However, within each implant location / animal, the prediction quality seems to be correlated with the amount of power in certain LFP frequency bands (0-10, 10-20 and 40-50 Hz, standardised LFPs). Secondly, the results suggest that bipolar LFPs are more informative as to SUA than unipolar LFPs. This suggests common mode rejection aids in the elimination of non-local neural information. Thirdly, the best individual bipolar LFPs generally perform better than when using all available unipolar LFPs. This suggests that LFP channel selection may be a simple but effective means of lossy data compression in Wireless Intracortical LFP-based BCIs. Overall, LFPs were moderately predictive of SUA, and improvements can likely be made.**

I. Background

Brain-Computer Interfaces (BCI) are devices for connecting nervous systems to electronics. The next generation of intracortical BCIs is expected to be wireless. This is to remove any transcutaneous connection and its associated infection risks. An informative signal with low sampling and communication rate is desirable for Wireless Intracortical BCIs (WI-BCI) to reduce power consumption. This is to limit heat dissipation into the brain [1], and to generally reduce design constraints. Since the advent of intracortical BCIs, four major representations of intracortical signals have gained prominence: Broadband recordings, the Local Field Potential (LFP) signal, Single-Unit Activity (SUA), and Multi-Unit Activity (MUA) [2]. Of these four, two may be of particular interest: LFP and SUA.

A. Local Field Potentials

The LFP consists of the large-amplitude low-frequency oscillation observed in the Broadband neural signal. It is typically obtained by lowpass filtering the Broadband signal at 100 to 300 Hz. It therefore has a low sampling rate, often in the hundreds of Hz, making it compatible with WI-BCIs.

In the motor cortex, LFPs have been found to contain information relative to motor function [3] [4] [5]. Additionally, high-gamma frequency LFPs have been found to be phase-locked to neural firing rates [6] [7]. LFPs are also robust to electrode micro-motions, and are chronically recordable [5].

The neural origins of LFPs are still poorly understood [8]. However, LFPs are believed to result from the sum of neural activity from cells within 0.5-5 mm laterally of the electrode, and centimeters vertically. This means that adjacent electrodes often contain redundant information.

This redundant information can be lessened with bipolar recordings, where the electrode is referenced to a nearby electrode. In contrast, unipolar recordings reference electrodes to a distant electrode with 'neutral' voltage. In bipolar recordings, the portion of the LFP that is due to distant neural activity is suppressed via common mode rejection. This allows the LFP contribution of only neurons in the vicinity of the electrode pair to be extracted. However, taking bipolar recordings from between n electrodes increases the number of recordings by $(n^2\text{-}n)/2$, much of which contains redundant information. As such, bipolar recording can be incompatible with bandwidth reduction. This is unless some channel selection strategy is used, discussed in Section I C.

B. Single-Unit Activity

SUA consists of identifying which neurons are firing and when from the broadband neural recording. Firstly, the raw broadband signal is highpass filtered to remove the LFP. Any observed spikes, generally with significant high-frequency components, are then clustered together based on shape and amplitude. SUA recordings are generally held to have the highest information content of the three major derivative representations (LFP, SUA, MUA). This is because the contributions of individual neurons to the overall signal are explicitly identified [2]. However, the SUA encoding often deteriorates in quality over the course of a few months after implantation of the electrodes. This is due to the foreign body response where the electrodes are encapsulated by fibrous scar tissue [9], which acts as a highpass filter, filtering out the spike shapes. Additionally, SUA may be prohibitively computationally expensive to compute on-implant in WI-BCIs without a computer-to-implant downlink [10]. As such, SUA is highly informative, but unstable and perhaps too computationally expensive for WI-BCIs.

C. Using LFPs to predict SUA

The correlation and relationship between SUA and LFP have been studied extensively [11]. SUA has even been predicted using $<$ 5 Hz LFPs with linear decoders [3]. Predicting SUA from LFPs is interesting because it indicates common information. A large amount of common

Oscar W. Savolainen is supported by an EPSRC DTP scholarship.

The authors are with the Department of Electrical and Electronic Engineering and the Centre for Bio-Inspired Technology, Imperial College London, SW7 2AZ, UK. Email: {o.savolainen18, t.constandinou}@imperial.ac.uk

information would suggest that a low-frequency, chronically available encoding of SUA would be available in the LFP. This would mean the easy-to-measure and stable LFP could be used as a WI-BCI compatible proxy for SUA, and may have neuroscience implications for the analysis of SUA.

However, to the best of the author's knowledge, no one has used Long Short-Term Memory (LSTM) decoders to predict SUA from LFPs. LSTMs are a form of Artificial Neural Network (ANN). They are one of the only Machine Learning algorithms that can perform non-linear sequence-to-sequence prediction [12]. As such, they are a very useful tool for predicting SUA from LFPs, where the relationship is presumed to be non-linear. Accordingly, LSTMs have performed very well in BCI Behavioral decoding tasks relative to other decoders [13] [14] [15].

As aforementioned, there can be significant redundancy across LFP channels. Classic feature reduction methods such as Principal Component Analysis (PCA) may be too computationally expensive in WI-BMIs. This is due to the significant continuous operations. Therefore, it would also be interesting to determine how well individual channels predict SUA relative to when using all available channels. If they perform well, it would indicate that effective channel selection could reduce bandwidth. As such, a very computationally inexpensive proxy for SUA would be available.

Therefore, this paper investigates whether LSTMs can be used to predict SUA from LFPs. Both unipolar and bipolar LFPs are considered. Additionally, this paper compares the effects of using single versus multiple LFP channels in the prediction. To the best of the authors' knowledge, neither of these points have been considered before.

II. Methods

A. Data

All work was done in MATLAB R2018a, using neural recording data from Andrew Jackson and Thomas Hall [3]. The data consists of raw Broadband microwire recordings from Macaque Primary Motor cortex, sampled at 16-bit resolution and 24.4 kHz. The data came from 3 animals (Dusty, River, Silver) and a 'Ukiah' dataset, which was understood to originate from one of the 3 animals but using a different microwire array. The recording lengths belong to vector RL = [205, 364, 396, 473, 646] s. Respectively, the values are the minimum recording length, the 1st quartile, the median, the 3rd quartile, and the maximum.

The SUA was extracted using Wave_Clus with standard settings [16]. MUA channels were not considered. The LFP was obtained by lowpass filtering the raw recording at 100 Hz with the `lowpass` function, which applies a 3rd order FIR filter. The LFPs were then downsampled to 500 Hz, and the SUA was binned using non-overlapping 2 ms bins. Bipolar LFPs were obtained by subtracting Unipolar LFPs from each other after filtering and downsampling.

B. Decoder structure and hyperparameters

The LSTM decoder consisted, in order, of a Sequence Input layer, an LSTM layer with 3 hidden units, a Fully Connected layer with number of neurons equal to the number of SUA channels, and a Regression layer. The Adam optimization algorithm was used, along with an initial learn rate of 0.005. A learn drop rate of a quarter of the number of epochs, and learn drop rate factor of 0.2 were used. The gradient threshold was set to 1.

C. Decoder parameter optimization

Firstly, the data was standardized. Then, various parameters were optimised for both the unipolar and bipolar LFP-SUA predictions. Firstly, whether to feed the data in as a continuous sequence or in overlapping segments of 2 s duration, inspired from the work in [3]. Secondly, if and how many times to smooth the binned SUA data using a moving average filter. Thirdly, the width of the potential moving average window. Lastly, the use of bidirectional v. unidirectional LSTM networks. No other LFP processing or feature extraction was considered other than the standardisation and potential segmentation.

A 75-25% training-testing split was used, along with a 75-25% training-validation split of the training data within the parameter optimisation. 1000 epochs were used for the continuous data and 100 for the segmented data. The parameter optimization was repeated 10 times for each parameter and recording combination, and the average result used.

The mean Pearson correlation value r between the predicted and actual SUA (averaged across all SUA channels) was used to measure the decoder performance.

D. Decoder evaluation

After parameter optimization, the final performance of the LSTMs was evaluated using the testing data. They were evaluated in two configurations. Firstly, they were tested in Multiple-Input-Multiple-Output (MIMO) configuration, where all of the LFP channels were used to predict the SUA from all of the electrodes. Secondly, they were tested in Single-Input-Multiple-Output (SIMO) configuration, where each individual LFP channel was used to predict the SUA from all of the electrodes. SIMO was used to determine the potential performance of channel selection. The same parameters were used for SIMO as MIMO. This is because the SIMO parameter optimization would have taken a prohibitive amount of time on the order of many months. The MIMO parameters were assumed to generalise to SIMO, as they mainly concerns SUA processing. The same data is contained in MIMO as in SIMO, if in aggregate form.

The results were obtained by taking the median performance of 5 identical networks trained for each recording and channel combination. Out of 66 available recordings, 7 failed due to time-out errors. The computational cluster used for the network training and testing limited job length to 24 h, which was exceeded by the failed jobs. As such, an n of 59 recordings is considered in the following analysis, with no other screening. The 7 failed recordings were not repeated due to timing constraints and because an n of 59 gave statistically significant results.

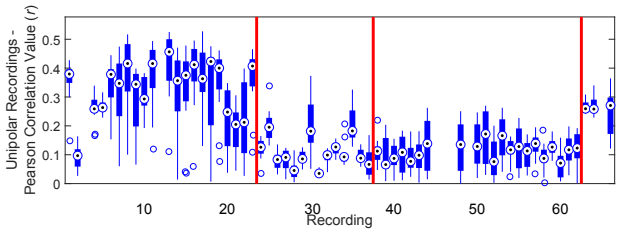

Fig. 1. Unipolar SIMO box plots. Each column represents a recording. The y-axis represents the box plot of each recordings' channels ability to predict the SUA. 12 electrodes were used in recordings 1-37, 24 in recordings 38-62, and 10 in recordings 63-64. Recordings 1-23 are from macaque Dusty, 24-37 from macaque River, 38-48 from macaque Silver, 49-62 from Silver using another setup, and 63-66 from 'Ukiah' recordings, subject unknown. Vertical red lines delineate animals. Empty columns represent failed recordings.

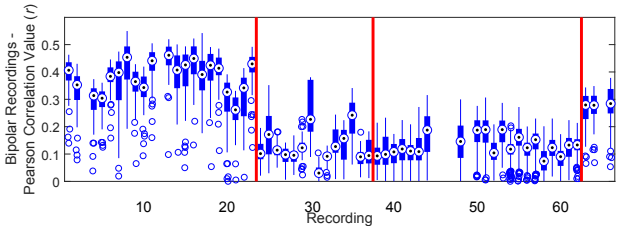

Fig. 2. Bipolar SIMO box plot. The dataset details are the same as for the Unipolar recordings in Fig. 1.

III. Results

A. Parameter optimisation results

The parameter optimization findings were that using continuous data, twice smoothing the SUA recordings with 0.5 s moving average window, and using a bidirectional LSTM gave the best performance. The bidirectionality increased r by approximately 0.03 across the different datasets and parameter combinations. In this paper bidirectional LSTMs were used. However, in future practical BMI-related work, unidirectional LSTMs should be used due to the real-time decodability constraint. 300 epochs were used for training. This is because the performances during network training were relatively stable, as observed graphically, while relatively rapid. Increasing the number of epochs lead to logarithmically better performance, at the cost of computation time.

B. Decoding test results

Table I contains the decoder test results. The Best Channel SIMO corresponds to the best performing LFP channel in each recording. It is the main channel that would be selected in a successful on-implant channel selection algorithm. The best channels were not necessarily the same channels for each recording, and were selected on a recording-by-recording basis. The results from using the median performing channel per recording are also included, i.e. Median Channel SIMO. This is because it is a channel that is likely to be selected in the case of an indiscriminatory channel selection. The SIMO channel box plots are shown in Fig. 1 and 2. For unipolar and bipolar modes respectively, they give the SIMO r between the actual and predicted SUA.

C. Findings

There were three major findings.

TABLE I

Pearson Correlation Between Actual and Predicted SUA for Specified Recording, Unipolar v. Bipolar LFPs and SIMO v. MIMO LFP Input

Recording		Unipolar channels			Bipolar channels		
		MIMO	Best Channel SIMO	Median Channel SIMO	MIMO	Best Channel SIMO	Median Channel SIMO
Recordings from All animals	Worst	9.6	5.6	3.5	6.9	10.4	3.1
	1st quartile	18.0	17.3	9.7	15.6	21.0	11.1
	Median	26.6	25.1	13.9	23.3	30.0	18.7
	3rd quartile	42.7	39.5	28.6	40.2	42.6	35.0
	Best	59.6	52.7	45.7	57.7	55.0	46.1
	Mean	30.1	28.4	19.8	27.9	31.7	22.7
Recordings from Dusty and 'Ukiah'	Worst	24.5	16.4	9.7	23.6	33.1	26.1
	1st quartile	39.2	34.4	25.9	37.6	38.5	31.9
	Median	46.1	43.6	35.2	44.4	45.4	38.1
	3rd quartile	49.1	47.7	40.3	48.7	49.7	42.5
	Best	59.6	52.7	45.7	57.7	55.0	46.1
	Mean	44.1	41.3	32.5	42.8	44.6	37.2
Recordings from Silver and River	Worst	9.6	5.7	3.5	6.9	10.4	3.1
	1st quartile	16.5	15.7	8.6	11.5	18.6	9.5
	Median	19.3	18.3	11.2	16.5	22.4	11.8
	3rd quartile	24.2	22.6	12.9	20.5	27.2	15.5
	Best	36.8	37.3	19.5	36.0	38.1	24.2
	Mean	20.1	19.5	11.1	17.2	22.9	12.8

One-sided t-test results:
(a) Bipolar Median Channel SIMO > Unipolar Median Channel SIMO
Dusty and 'Ukiah': p = 1.6e-4; Silver and River: p = 6.9e-5
(b) Bipolar Best Channel SIMO > Unipolar Best Channel SIMO
Dusty and 'Ukiah': p = 3e-3; Silver and River: p = 3.4e-10
(c) Bipolar Best Channel SIMO > Unipolar MIMO
Dusty and 'Ukiah': p = 0.294; Silver and River: p = 1.9e-7

Pearson correlation values given in %.

1) Significant unexplained variation between animals: The Unipolar Best Channel SIMO for every Dusty and 'Ukiah' recording produced a median r of 43.6%, and the bipolar equivalent was 45.4%. Together, the River and Silver recordings only had a best-channel unipolar median of 18.3%, and a bipolar median of 22.4%.

Furthermore, Linear Regression Analysis using `fitlm` was performed. The Unipolar MIMO performance was used as an estimate for the recording's prediction quality. The mean standard deviation, pre-standardisation but post-smoothing, of the SUA was used as a measure of the neurons' level of activity. Also considered was the recording-averaged LFP power across the channels in 10 Hz bands, with 5 Hz overlap, between 0 and 100 Hz. The power was taken from standardized LFPs. The recording length was also included. The Unipolar MIMO performance had a bimodal distribution, with Dusty and 'Ukiah' recordings performing better than River and Silver recordings. As such, a categorical variable $c \in [0\ 1]$ was used to delineate Dusty and 'Ukiah' recordings (1) from Silver and River (0). This led to approximately normal distributions within each category.

During the analysis, predictors were sequentially removed from the model if they had p-values below 0.05, or if their inclusion did not increase the adjusted R^2 value by more than 0.01. The final result of the regression analysis is given in Table II. The model's adjusted R^2 is 0.824 and has a p-value

TABLE II
LINEAR REGRESSION ANALYSIS RESULTS

Variable	Type	Estimate	tStat	p-value
Intercept	Continuous	-0.0966	-1.11	0.271
0-10 Hz	Continuous	0.305	2.87	0.00584
10-20 Hz	Continuous	0.717	3.32	0.00162
40-50 Hz	Continuous	1.08	2.54	0.0138
Dusty and Ukiah	Categorical	0.214	11	2.11E-15

of 1.5e-20. The Durbin-Watson test does not show significant 1st order autocorrelation (p=0.12). Removing the categorical variable c reduces the adjusted R^2 value to 0.44. Removing the LFP power variables reduces the adjusted R^2 value to 0.75. As such, it seems that there is significant variability between animals that is unexplained. However, within each category of animal, the mean power in certain LFP frequency bands is a statistically significant indicator of the quality of the MIMO LFP-to-SUA prediction.

2) Bipolar channels seem to be broadly more informative than unipolar channels: In SIMO, bipolar recordings consistently outperform unipolar recordings. This is true, across animals, whether the best or the median performing channel per recording is used. This can be observed in Table I; One-sided t-test results (a) and (b). Interestingly however, in MIMO, unipolar channels seem to outperform bipolar channels. This is even though they contain the same data. However, in unipolar MIMOs, the data is in a condensed form. Additionally, the bipolar Best Channel SIMOs mostly outperform the bipolar MIMOs. MIMOs have access to all the data used in the Best Channel SIMOs, and more. These two observations suggest that the Curse of Dimensionality (CoD) [17] is likely a factor in bipolar MIMOs.

3) Significant reductions in channel count, and thus bandwidth, seem to be possible with effective channel selection: This is evidenced by the Bipolar Best Channel SIMO $>$ Unipolar MIMO result. This can be observed in Table I; One-sided t-test result (c). Reducing the bandwidth via effective channel selection does not result in significant losses in SUA predictivity. Here, effective channel selection means selecting the Bipolar Best Channel SIMO. In Silver and River, effective channel selection even seems to results with gains in LFP-to-SUA decoding performance. In Dusty and 'Ukiah' the strictly greater result was not statistically significant. The bandwidth savings varied from 90% to 96% depending on the original number of electrodes (10 to 24).

IV. DISCUSSION

A. Improving LFP-SUA prediction

Overall, the results are mixed. On the one hand, the best MIMO case established a 59.6% correlation between actual and predicted SUA, across recordings hundreds of seconds in length with limited optimisation and no LFP feature extraction. On the other hand, even in the best case, $1 - 0.596^2 = 64.5\%$ of the variance in the observed SUA has not been explained. As such, it is unclear from this paper whether LFPs can be used as a proxy of sufficient quality for SUA. It could be interesting to see to what extent using only the top 2-3 channels improves decoder performance. The CoD may have been a factor even in the unipolar MIMO decoders. As such, one could limit the number of channels in MIMO configuration to 2-3 custom-selected channels, especially bipolar channels. This may improve the decoder performances. Increasing the number of epochs beyond 300 is also likely to increase performance.

B. LFP feature extraction

No feature extraction was implemented on the LFP data. Since the desirable features are unknown, it was left to the decoder to apply non-linear transformations to the data. Future work should be concerned with the identification of these features.

V. ACKNOWLEDGEMENTS

We thank Andrew Jackson and Thomas Hall for making their data [3] available.

REFERENCES

[1] P. D. Wolf and W. Reichert, "Thermal considerations for the design of an implanted cortical brain–machine interface (BMI)," *Indwelling Neural Implants: Strategies for Contending with the In Vivo Environment*, pp. 33–38, 2008.

[2] M. A. Lebedev and M. A. Nicolelis, "Brain-machine interfaces: From basic science to neuroprostheses and neurorehabilitation," *Physiological reviews*, vol. 97, no. 2, pp. 767–837, 2017.

[3] T. M. Hall *et al.*, "Real-time estimation and biofeedback of single-neuron firing rates using local field potentials," *Nature communications*, vol. 5, p. 5462, 2014.

[4] T. Milekovic *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.

[5] R. D. Flint *et al.*, "Long term, stable brain machine interface performance using local field potentials and multiunit spikes," *Journal of neural engineering*, vol. 10, no. 5, p. 056005, 2013.

[6] S. Ray *et al.*, "Neural correlates of high-gamma oscillations (60–200 Hz) in macaque local field potentials and their potential implications in electrocorticography," *Journal of Neuroscience*, vol. 28, no. 45, pp. 11 526–11 536, 2008.

[7] G. Buzsáki and E. W. Schomburg, "What does gamma coherence tell us about inter-regional neural communication?" *Nature neuroscience*, vol. 18, no. 4, p. 484, 2015.

[8] O. Herreras, "Local field potentials: myths and misunderstandings," *Frontiers in neural circuits*, vol. 10, p. 101, 2016.

[9] K. M. Szostak *et al.*, "Neural interfaces for intracortical recording: Requirements, fabrication methods, and characteristics," *Frontiers in Neuroscience*, vol. 11, p. 665, 2017.

[10] S. Luan *et al.*, "Compact standalone platform for neural recording with real-time spike sorting and data logging," *Journal of neural engineering*, vol. 15, no. 4, p. 046014, 2018.

[11] M. Rasch *et al.*, "From neurons to circuits: linear estimation of local field potentials," *Journal of Neuroscience*, vol. 29, no. 44, pp. 13 785–13 796, 2009.

[12] C. Olah, "Understanding LSTM networks," 2015.

[13] N. Ahmadi *et al.*, "Decoding hand kinematics from local field potentials using long short-term memory (LSTM) network," 2019, pp. 1–5. [Online]: http://arxiv.org/abs/1901.00708v1

[14] J. I. Glaser *et al.*, "Machine learning for neural decoding," *arXiv preprint arXiv:1708.00909*, 2017.

[15] T. Hosman *et al.*, "BCI decoder performance comparison of an LSTM recurrent neural network and a kalman filter in retrospective simulation," in *2019 9th International IEEE/EMBS Conference on Neural Engineering (NER)*. IEEE, 2019, pp. 1066–1071.

[16] R. Q. Quiroga *et al.*, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering," *Neural computation*, vol. 16, no. 8, pp. 1661–1687, 2004.

[17] F. Y. Kuo and I. H. Sloan, "Lifting the curse of dimensionality," *Notices of the AMS*, vol. 52, no. 11, pp. 1320–1328, 2005.