Cortical inputs underlying Beta Event Amplitude Modulation

Nicholas Tolley[1,2], Stephanie R. Jones[1]
Department of Neuroscience[1], Data Science Initiative[2] (Brown University)

Introduction

Magneto- and electroencephalography (MEG/EEG) are non-invasive neural recording techniques that primarily reflect activity from the neocortex, the outermost surface of the brain. One prominent activity pattern observed in MEG/EEG is that of Beta Events (BE), a transient stereotyped waveform that exhibits spectral power in the 15-29 Hz frequency range. BE have been associated with an extensive range of sensorimotor functions, but it's circuit level origins and functional significance remain undercharacterized (Kilavik et al., 2013; Sherman et al., 2016; Shin et al., 2017).

An emerging approach to elucidating the origins of macroscale neural signals is the simulation of biological neural networks. This proposal utilizes the Human Neocortical Neurosolver (HNN), a biophysically detailed model of a single cortical column that accurately reproduces MEG/EEG activity from circuit level spiking evoked by proximal and distal extrinsic inputs (**Figure 1A**).

Recent literature has demonstrated that the amplitude of BE exhibits significant age related changes (Brady et al., 2020). However, there are currently no predictions on the circuit level mechanisms underlying amplitude modulation. To address these problems, this study 1) employs masked autoregressive flows to capture the distribution of parameters producing BE (Papamakarios et al., 2018), and 2) generates a simulated dataset of BE with the goal of predicting the amplitude of the waveform from simulation parameters using regression analysis.

Proximal and distal inputs were parameterized by 8 values (**Figure 1B)**. Both proximal and distal inputs were assigned parameters for the mean (ms) and variance (ms^2) of input spikes. Two additional parameters were included to modulate the strength of the inputs to excitatory (pyramidal) and inhibitory (basket) neurons. Strength is in units of nano-siemens (nS) which quantifies the maximum electrical conductance at a synapse during neurotransmission. The target variable, BE amplitude, is measured in nanoampere meters (nAm).

Exploratory Data Analysis

BE are characterized by a stereotyped triphasic waveform, with a dominant negative peak whose amplitude is likely biologically significant. This project aims to uncover how model parameters defining extrinsic inputs to the cortex can be used to predict BE amplitude. 100,000 samples were generated from a fitted MAF which captures the distribution of BE producing parameters. Examples of large and small amplitude BE are shown in **Figure 2A**. The distribution of BE Amplitudes is unimodal with a peak at -5000 nAm (**Figure 2B**). Amplitudes generally range from [-6000, -1000] nAm.

Figure 3 visualizes the distribution of parameter samples. Parameters defining the mean and variance of extrinsic drives tend to exhibit low variance unimodal distributions. Synaptic weight parameters however exhibit higher variance univariate distributions, with complex non-linear interactions visible in the bivariate plots.

Exploratory analysis revealed a potential interaction between distal variance and distal mean time with respect to BE Amplitude. **Figure 4** visualizes this interaction as a scatter plot with distal variance and mean time on the x and y axes respectively, and a color map indicating BE amplitude. A line of high amplitude BE is visible in the center of the distribution, with a slight positive correlation suggesting a cooperative mechanism underlying amplitude modulation.

Methods

<u>Preprocessing and cross validation</u>

The dataset was generated by sampling from a masked autoregressive flow (MAF) model trained to capture the BE producing region of the parameter space. Since all observations were sampled from the same underlying distribution, the dataset is independent and identically distributed (IID). The parameters used as features to predict BE amplitude are all continuous with no obvious min/max boundary, as such a standard scaler was used to transform the data prior to model training.

The data set was split using a standard 60/20/20 splitting strategy for the train, test, and validation sets respectively. The standard scaler was fit to the training set features, and then used to transform the test and validations sets.

<u>Machine Learning Algorithms Compared</u>

Multiple regression algorithms were compared for their ability to predict BE amplitude from parameters of extrinsic inputs. R2 score was used as a comparison metric. Specifically, algorithms were selected from the scikit-learn python toolbox and included LinearRegression, Ridge, and KNeighborsRegressor. We additionally employed the XGBoost regressor.

Linear regression was chosen for its simplicity and interpatibility. No parameters are necessary for tuning. Ridge net was chosen for its ability to penalize coefficient magnitude. Parameter sweeps for the L2 regularization ('alpha') parameter included 10 logarithmically spaced values on [10e-2, 10e2]. KNeighborsRegression parameter tuning included the 'n_neighbors' parameter with 10 linearly spaced values on [1,100]. XGBoostRegressor was chosen due to well recognized high performance on a wide range of machine learning problems. The max depth parameter was tuned on the values [1,2,3,5,10,20,30].

Two sources of stochasticity were controlled in this analysis. The first is stochasticity in the training algorithm itself, where a random state is used to initialize the model (only Ridge and XBoost). The second form of stochasticity comes from the random partitioning of the train/test/validation sets. To control for both forms of stochasticity, model training, hyperparameter tuning, and testing were performed separately for 5 different random states.

Model parameters for reported results were chosen based on the best scores on the validation set. To ensure that the parameter ranges were sufficiently large, training and validation scores were stored and inspected to ensure both overfitting and underfitting.

To compare the performance of each model, test scores were averaged for the best performing model across all random states. A baseline test score was calculated by using a constant value prediction function set to the mean BE amplitude of the training set.

Results

Among the regression models trained to predict BE amplitude, the best performance came from XGBoost, followed closely by KNeighbors. **Figure 5** plots true vs. predicted BE amplitudes for the test set on all models. It is immediately apparent that linear and ridge regression perform worse on observations with either small or large amplitudes relative to the mean. Model performance was further quantified by the R2 score on the test set (averaged across 5 random initialization, see methods). **Figure 5** Demonstrates that XGBoost is indeed the best performing algorithm with a test score of 0.680 ± 0.005.

To better understand what features models were using to predict BE amplitude in terms of global feature importance, the beta coefficients of linear and ridge regression were inspected, as well as the weight and gain coefficients of the XGBoost model (**Figure 6**). Linear and ridge regression both indicate that proximal mean time, and proximal input strength to layer 5 neurons, are the most important predictors of BE amplitude. Similarly, the gain coefficient of the XGBoost model supports proximal mean time as an important predictor.

Permutation tests were performed as an alternative global feature assessment analysis to see how test scores of each model decrease after shuffling of individual features (**Figure 7).** Both XGBoost and KNeighbors suggest that proximal mean time and variance are the most impactful on test scores.

Finally, local feature importance of the XGBoost model was assessed via SHAP scores (**Figure 8**). Multiple parameters exhibited a critical point-like behavior where the additive contribution to BE amplitude flipped from negative to positive along the range of the parameter values.

It is surprising that proximal input parameters were consistently strong predictors of BE amplitude across multiple models and feature importance metrics. **Figure 1** demonstrates how the immediate effect of proximal input leads to an increased voltage which would intuitively decrease BE amplitude. However, these results suggest that the dynamics of network level spiking activity produce a more complex relationship such that the proximal inputs have a strong influence over BE amplitude.

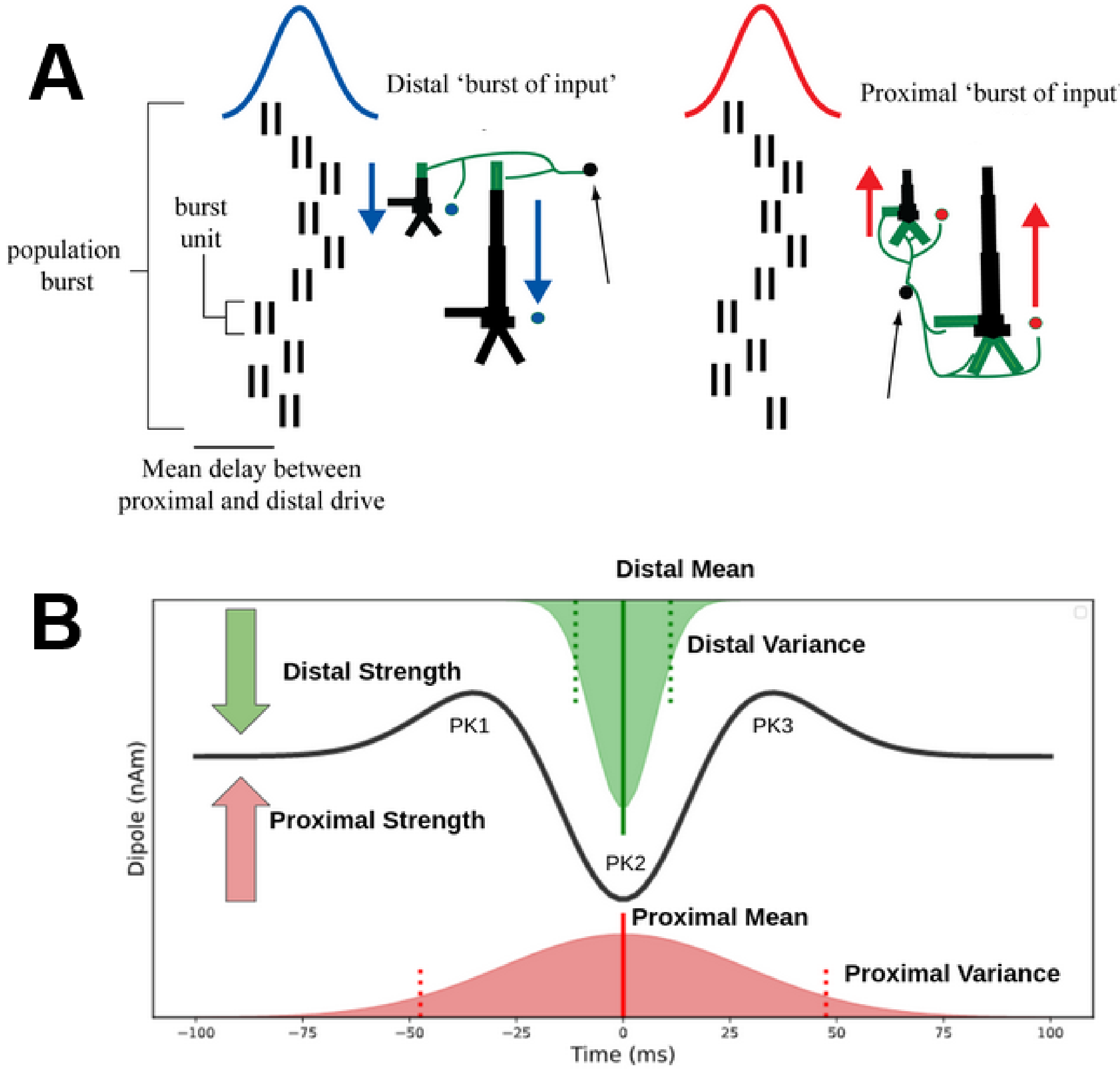
Outlook

While prediction performance was high for the XGBoost model, several improvements can be made for future studies. As shown in the EDA, this distribution of BE amplitudes is non-uniform. Predictions for small or large amplitude events may be worse since the model sees mostly intermediate size BE amplitudes in training/testing. One future approach may be to threshold the dataset and train separate models on small and large amplitude events to investigate if feature importances change in the two different regimes.

Another improvement is to engineer new features. The parameters were selected due to the interface used to control simulations. Newly constructed features using domain expertise may better predict BE amplitude. Additionally for data collection, more simulation parameters could be allowed to vary, which may help improve the biological interpretability of model results, at the cost of more complex relationships between features. Finally, further improvements to XGBoost predictions may be possible with more rigorous hyperparameter tuning with features besides max depth.

Code Availability

The command line interface version of the Human Neocortical Neurosolver software: https://github.com/jonescompneurolab/hnn-core
Code used to preprocess and generate the figures included in this report:
https://github.com/ntolley/data1030_project



<u>Figure 1: Parameters of extrinsic inputs influence the shape of Beta Events</u>
The parameters controlling the generation of be are shown.
A: Exogenous proximal and distal inputs to the cortical column are produced by sampling spike times from a gaussian distribution.
B: For both proximal (red) and distal (green) drives, the mean, standard deviation, and strength determine the characteristics of gaussian distributed input spikes.

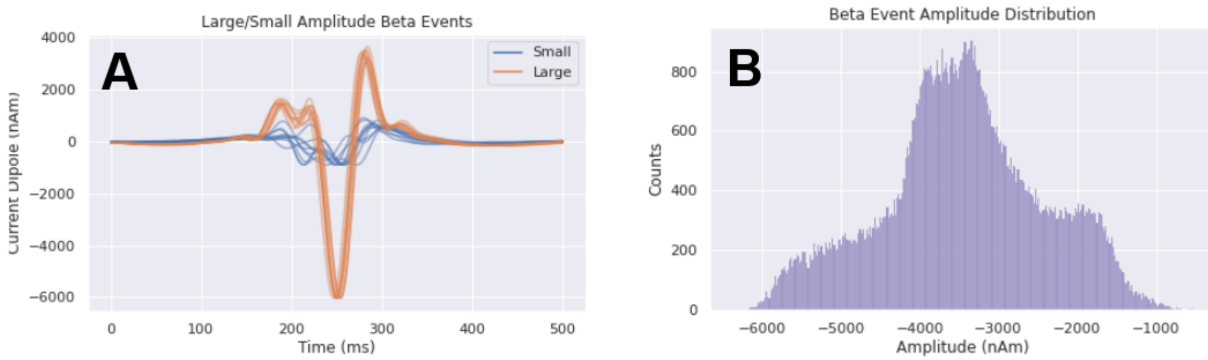


Figure 2: Distribution of BE amplitudes

A: Exemplar simulations of small (blue) and large (orange) BE. **B:** 100,000 BE were simulated, the distribution of amplitudes is shown with a primarily unimodal with a peak centered near -5000 nAm.

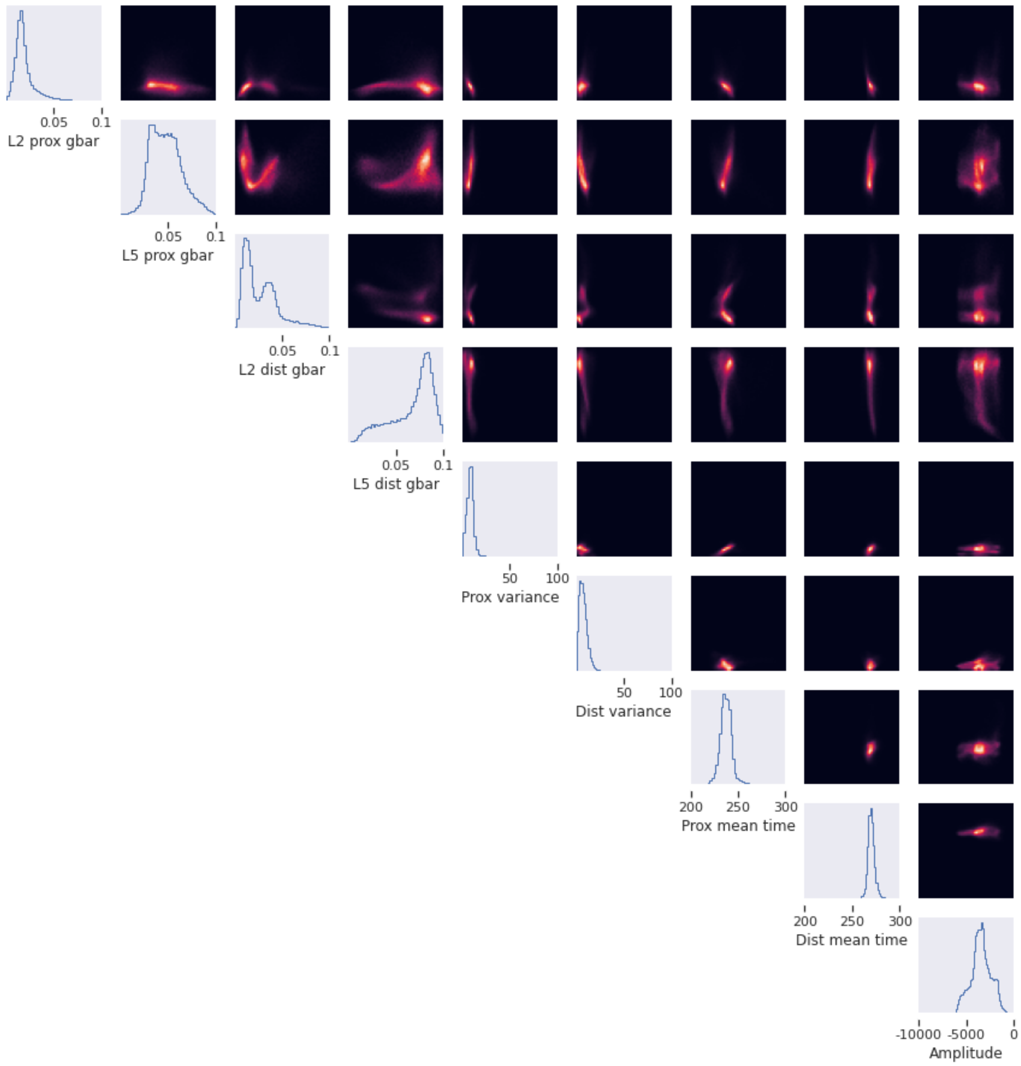


Figure 3: Masked autoregressive flow captures Beta Event producing parameter subspace

Masked autoregressive flows (MAF) were used to approximate the distribution of the BE producing subspace of the parameters. 100,000 samples were generated from the MAF whose distributions are plotted above. Univariate marginal distributions are plotted on the diagonal, and bivariate marginals are plotted on the off diagonal squares.

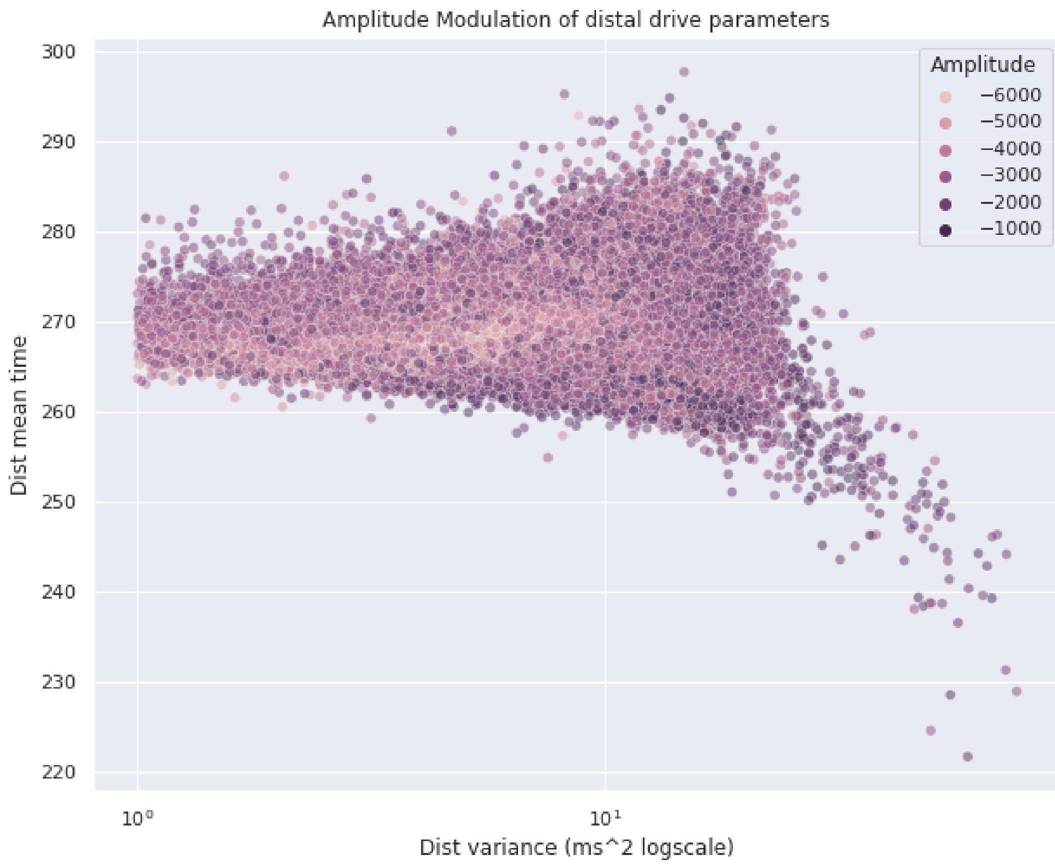
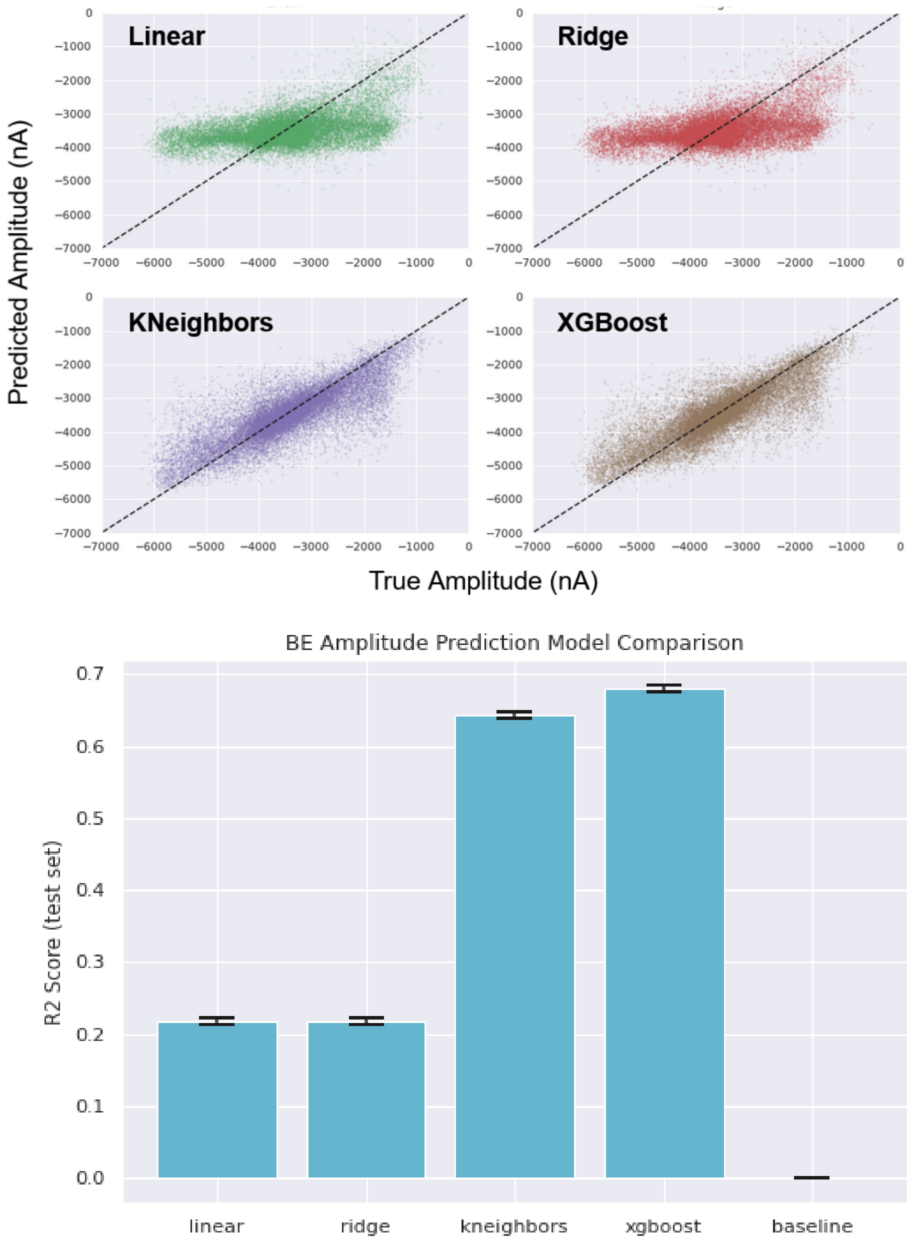


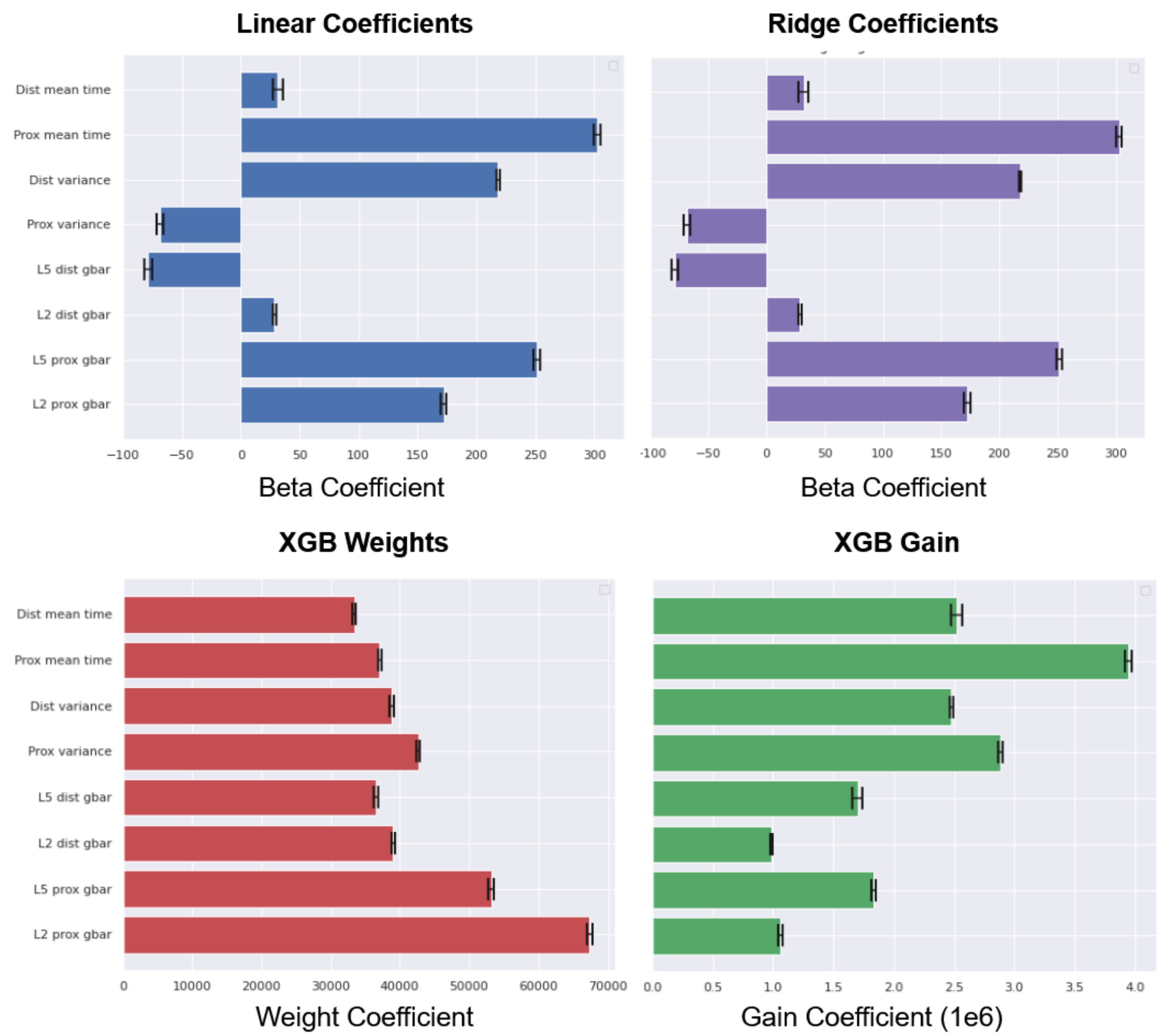
Figure 4: Distal drive parameters interact to modulate Beta Event Amplitude

‘Dist variance’ and ‘Dist mean time’ parameters are plotted with a color scale corresponding to the amplitude of the BE produced by simulating these parameters.



<u>Figure 5: BE Amplitude Prediction Performance</u>

True and predicted BE amplitude for the test set of each mode is shown (top). Data points lying on the dotted black line indicate perfect regression performance. As shown, linear and ridge regression qualitatively perform poorly. Prediction performance quantified by R2 scores (bottom) indicate that XGBoost produces the highest accuracy on the test set, only slightly above KNeighbors.



<u>Figure 6: Model coefficients reveal importance of proximal parameters</u>
Beta coefficients for linear and ridge regression models (top) indicate that proximal mean time, and proximal strength to layer 5 neurons (L5 prox gbar) are the strongest predictors of BE amplitude. In contrast, XGBoost weights (bottom left) suggest that proximal strength to layer 2 neurons (L2 prox gbar) is more important. The XGB gain (bottom right) coefficient coincides with linear models and indicates proximal mean time as the most predictive of BE amplitude.

Figure 7: Proximal mean time and variance are most important for test scores

Permutation testing was used to compare the impact of each feature on XGBoost and KNeighbors models. The true test score of each model is shown by the blue line. As shown in both models, permuting proximal mean time and proximal variance parameters leads to the largest decrease in model test scores.

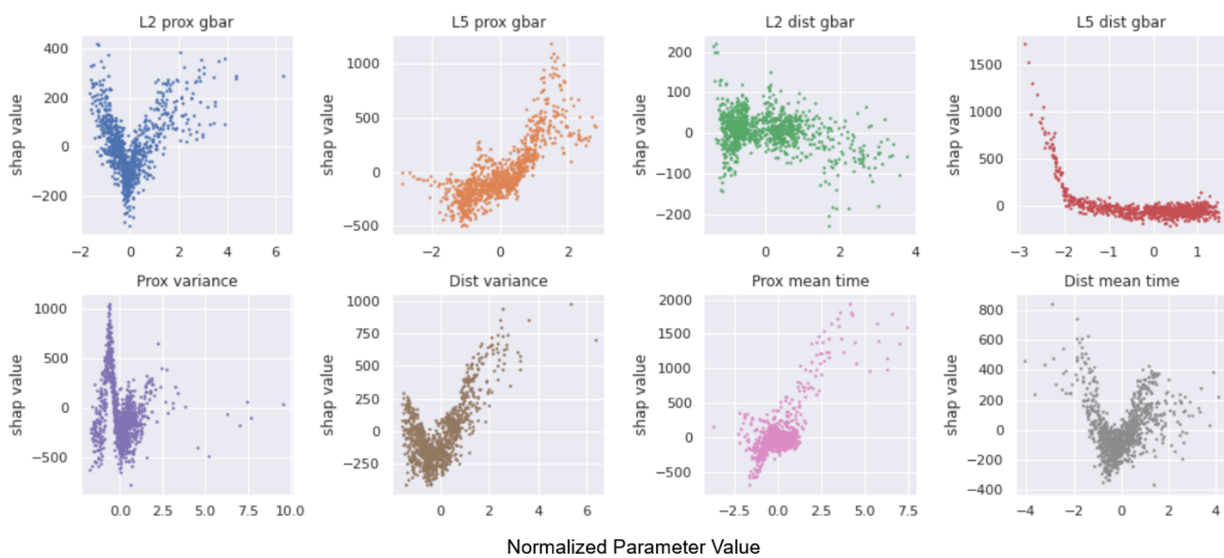


Figure 8: SHAP Scores indicate critical point in multiple parameters

SHAP scores were calculated for a test set with 1000 randomly selected points. Scores can be interpreted as additive contributions, therefore negative values indicate the feature predicts larger amplitude BE. As shown, multiple parameters exhibit a nonlinear relationship where negative and positive contributions to BE amplitude predictions flip, notably proximal/distal variance.

References

Bonaiuto, J. J., Little, S., Neymotin, S. A., Jones, S. R., Barnes, G. R., & Bestmann, S. (2021). Laminar dynamics of high amplitude beta bursts in human motor cortex. *NeuroImage*, 118479. https://doi.org/10.1016/j.neuroimage.2021.118479

Brady, B., Power, L., & Bardouille, T. (2020). Age-related trends in neuromagnetic transient beta burst characteristics during a sensorimotor task and rest in the Cam-CAN open-access dataset. *NeuroImage*, *222*, 117245. https://doi.org/10.1016/j.neuroimage.2020.117245

Kilavik, B. E., Zaepffel, M., Brovelli, A., MacKay, W. A., & Riehle, A. (2013). The ups and downs of beta oscillations in sensorimotor cortex. *Experimental Neurology*, *245*, 15–26. https://doi.org/10.1016/j.expneurol.2012.09.014

Law, R. G., Pugliese, S., Shin, H., Sliva, D., Lee, S., Neymotin, S., Moore, C., & Jones, S. R. (2021). Thalamocortical Mechanisms Regulating the Relationship between Transient Beta Events and Human Tactile Perception. *Cerebral Cortex*. https://doi.org/10.1093/cercor/bhab221

Papamakarios, G., Pavlakou, T., & Murray, I. (2018). Masked Autoregressive Flow for Density Estimation. *ArXiv:1705.07057 [Cs, Stat]*. http://arxiv.org/abs/1705.07057

Sherman, M. A., Lee, S., Law, R., Haegens, S., Thorn, C. A., Hämäläinen, M. S., Moore, C. I., & Jones, S. R. (2016). Neural mechanisms of transient neocortical beta rhythms: Converging evidence from humans, computational modeling, monkeys, and mice. *Proceedings of the National Academy of Sciences*, *113*(33), E4885–E4894. https://doi.org/10.1073/pnas.1604135113

Shin, H., Law, R., Tsutsui, S., Moore, C. I., & Jones, S. R. (2017). The rate of transient beta frequency events predicts behavior across tasks and species. *ELife*, *6*, e29086. https://doi.org/10.7554/eLife.29086

Williamson, S. J., Romani, G.-L., Kaufman, L., & Modena, I. (2013). *Biomagnetism: An Interdisciplinary Approach*. Springer Science & Business Media.