Learning to Learn: Modeling Neural Connectivity Dynamics During BCI-Driven Plasticity

William Ong
University of Washington

Abstract

Understanding the underlying neural connectivity dynamics during brain-computer interface (BCI) learning remains a central challenge in neuroscience research. In this study, we analyze the connectivity dynamics of neural activity data collected by the Allen Institute for Neural Dynamics involving the primary motor cortex of mice, featuring two-photon calcium imaging, a BCI task with a conditioned neuron (CN), and optogenetic photostimulation. Statistical analysis revealed a significant, global decrease in neuronal connectivity post-BCI and no detectable dependency between connectivity changes and spatial proximity. Using a series of random forest classifiers, we evaluated whether pre- and post-BCI connectivity patterns could be distinguished algorithmically based on different preprocessing and representation methods. Principal component analysis (PCA) on conditioned neuron connectivity vectors resulted in high classification accuracies, suggesting the presence of structured changes in the pre- and post-BCI CN connectivity. These results support the theory that learning modulates cortical connectivity in structured and interpretable ways, with implications for modeling synaptic credit assignment and enhancing BCIs.

1 Introduction

Neural populations are often conceptualized as dynamical systems, with their activity and connectivity patterns continually evolving in response to varying stimuli and feedback [1, 2]. Modeling the underlying dynamics of neuronal activity has been a central focus of neuroscience research, with significant advancements achieved by techniques such as two-photon calcium imaging, dimensionality reduction, and dynamical systems modeling [1,3,4]. Neural connectivity dynamics are typically inferred indirectly from neuronal activity and have received comparatively less attention, remaining a major open question in the field. Modeling these dynamics is particularly challenging due to their latent nature, high dimensionality, and the diverse nonlinear dynamics exhibited by individual neurons [5]. Distinguishing causal synaptic changes from neuron co-modulation presents an additional layer of complexity in interpreting connectivity dynamics. Nevertheless, a deeper understanding of neural connectivity dynamics, particularly neural plasticity, can unravel the complexities underlying neurological and psychiatric disorders, and advance brain-computer interfaces (BCIs) and neuroprosthetics through improved decoding capabilities [6].

In this study, we analyze and model neural connectivity dynamics using data generously provided by the Allen Institute for Neural Dynamics. Their recent experiments address the "credit assignment problem," which involves learning-related plasticity and how the brain modifies synapses and behavioral phenotypes to adapt to variable environments [7, 8]. They employ an optical brain-computer interface (BCI) task, where mice learn to control a reward port through the activity of a single conditioned neuron. Combined with two-photon (2P) calcium imaging and optogenetic stimulation performed before and after the BCI task, this approach enables tracking of causal connectivity changes in individual neurons before and after learning. We begin with a statistical analysis of the data to identify patterns indicative of neural plasticity and underlying connectivity dynamics, contextualizing our findings with prior research and theoretical models. Then,

we train multiple random forest classifiers using various input representations to identify and classify neural connectivity patterns observed during BCI learning. By analyzing population-level activity and causal stimulation data, we aim to uncover the principles of synaptic credit assignment and further the understanding of neural plasticity and connectivity dynamics.

2 Methods

2.1 Datasets

We use data collected by the Allen Institute for Neural Dynamics (AIND) during their credit assignment experiment, which provides insight into how cortical circuits of mice adapt to control a BCI using individual neurons in the motor cortex [7]. The experiments involved an optical BCI task, where mice learned to control a reward port through the activity of a single conditioned neuron (CN). Using two-photon (2P) calcium imaging, the AIND recorded neural activity of the head-fixed mice from layer 2/3 excitatory neurons in the primary motor cortex (M1) across multiple days. Each day, a new CN was selected, and its activity was mapped to the position of a motorized reward port in real-time. To move the port into reach and receive water rewards, mice had to increase the CN's activity to exceed a predefined threshold. Experimental sessions also included simultaneous photostimulation and calcium imaging sessions, where a single neuron or group of neurons was optogenetically stimulated to measure their causal influence on the neuronal network. Each experiment session had 5 experimental epochs: pre-BCI photostimulation, spontaneous activity, BCI behavioral task, spontaneous activity, and post-BCI photostimulation, allowing analysis of neural connectivity dynamics before and after learning. Connection mapping sessions were repeated daily, and imaging data was processed with Suite2p, which produced motion-corrected fluorescence traces, extracted regions of interest (ROIs), inferred spiking events, and stimulus-aligned behavioral metadata [9]. The data of each experimental session was packaged into a Neurodata Without Borders (NWB) format, containing information on the optogenetic stimulation trials, BCI task trial information, image segmentation with ROI classifications, and neural activity traces, which are motion-corrected fluorescence traces that define $\Delta F/F$, the change in fluorescence from baseline at each frame. We conduct our analyses and experiments on 25 individual experimental sessions, involving 6 unique subjects and experiments performed on different days. Each session recorded between 363 and 4465 neurons, with a median of 597 and a mean of 1885 neurons per session.

2.2 Statistical Analysis

To determine whether BCI learning influences neural connectivity, we calculated and compared correlation coefficients between fluorescence time series traces ($\Delta F/F$) recorded before and after the BCI task. These traces were extracted from the NWB files corresponding to multiple experimental sessions, identified by subject-specific identifiers and session timestamps. The conditioned neuron was localized by obtaining its coordinates from the BCI trial data, and the nearest neuron was identified by computing the Euclidean (2-norm) distance to each neuron's ROI coordinates. Pre- and post-conditioning periods were defined by the presence of "photostim" and "photostim_post" in the optogenetic stimulation data for the `stim_name` attribute, respectively. To quantify connectivity between the CN and other neurons, we computed pairwise Pearson correlation coefficients for all activity traces within the same period, then isolated the CN with all other neurons. For each experimental session, we calculated the mean and standard deviation of the correlation coefficients within each period to quantify the CN's functional connectivity with other neurons and the variability of those connection strengths. We then conducted a paired t-test comparing pre- and post-BCI connectivity distributions to evaluate BCI learning's influence on connectivity. We computed the Euclidean distances between the CN and other neurons based on their spatial centroids, and used Pearson's, Spearman's rank, and distance correlation to determine whether changes in connectivity were related to

spatial proximity. Furthermore, we compared the paired t-test results of 100 non-conditioned neurons to those of the conditioned neuron to determine whether the conditioning procedure produced statistically significant changes in neural connectivity. Finally, we aggregated results across all experimental sessions to identify group-level trends in connectivity modulation and to increase statistical power.

To further validate our findings, we quantified the number of neurons functionally connected to the CN before and after the BCI task by analyzing changes in their response to CN-targeted photostimulation. These CN-targeted photostim trials were identified by thresholding the 2-norm distance between the CN and photostim locations. The threshold was defined such that at least 50% of experimental sessions included one or more photostims within the radius of the CN, balancing computational overhead and detection accuracy (see Section 4 for more details). For each photostimulation event, a neuron's baseline activity was defined as the 1 second preceding stimulation, and response activity as the 1 second following stimulation. We then performed paired t-tests comparing the baseline and response activity across trials, where neurons exhibiting a statistically significant change in activity ($p < 0.05$) were classified as CN-connected.

2.3 Machine Learning Analysis

2.3.1 Pre- vs. Post-BCI Connectivity Classification

To determine whether BCI learning induces structured, model-detectable changes in neural connectivity, we trained random forest classifiers to differentiate between pre- and post-BCI neural activity segments. We selected the random forest classifier approach for its robustness to noise, ability to capture non-linear relationships, and effectiveness in modeling high-dimensional features spaces without relying on strict assumptions about data distributions [10]. Due to variability in neuron counts across experimental sessions, we evaluated four different preprocessing and representation methods for aggregating the data. In each method, we computed Pearson product-moment correlation coefficients of the motion-corrected fluorescence time series traces ($\Delta F/F$) for the pre- and post-BCI periods, which we defined as the connectivity matrices. In two of the methods, we extracted the upper triangular elements (excluding the diagonal) from the connectivity matrices to prevent redundancy and high-computation costs. Then, we trained two models: one using summary statistics (mean, standard deviation, 10th percentile, and 90th percentile) and another using principal component analysis (PCA). For PCA, we zero-padded the extracted elements to ensure consistent dimensionality and evaluated 10, 15, and 20 principal components. In the other two methods, we focused specifically on the CN connectivity within the connectivity matrices. Given the reduced data volume and computational demand, one method used additional statistical features, including the minimum, maximum, and the count of values that exceeded a threshold tuned by grid search. The other model applied PCA to the CN vectors, where we evaluated 10, 30, and 50 principal components. Each model used 100 decisions trees, was evaluated using 5-fold cross-validation, and all data were normalized to zero mean and unit variance.

2.3.2 Real vs. Synthetic CN Connectivity Change Classification

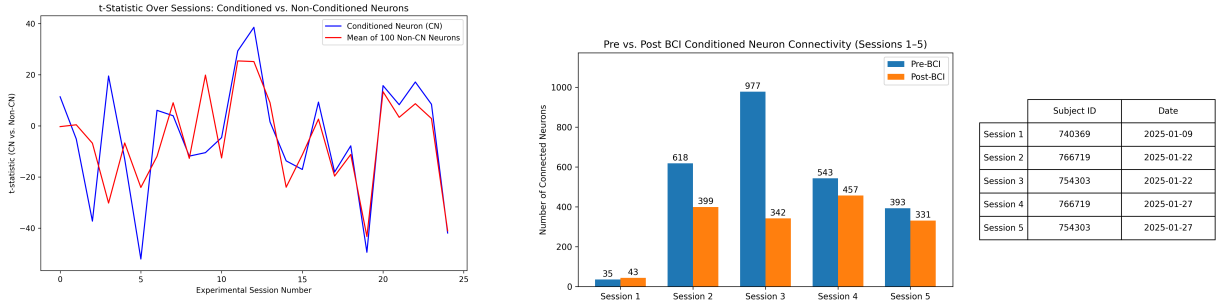
Beyond classifying pre- and post-BCI neural activity segments, we further examined whether within-session changes in CN connectivity exhibited systematic patterns or only reflected random variation. To assess this, we trained a random forest classifier to distinguish real from synthetic CN connectivity changes across the pre- and post-BCI transition. Consistent with the previous models, we defined CN connectivity as the CN-isolated vectors of Pearson product-moment correlation coefficients of $\Delta F/F$. Real changes were computed as the difference between pre- and post-BCI connectivity within the same session, capturing session-specific neural dynamics potentially associated with BCI learning and plasticity. In contrast, synthetic changes were generated by subtracting the pre-BCI connectivity of one session from the post-BCI connectivity of a different session, simulating variability across unrelated sessions. In both cases, we computed the differences by subtracting the extracted statistical features, which included a variable threshold tuned by grid search.

With this design, classifier performance exceeding chance would indicate that within-session CN connectivity changes reflect structured patterns rather than random noise. The model used the same configuration, evaluation metrics, and normalization procedure as in the previous classification task.

3 Results

3.1 Statistical Analysis

Across 25 sessions, the conditioned neuron exhibited a mean correlation coefficient t-statistic of -4.53 between pre- and post-BCI connectivity, with an associated mean p-value of 0.00477. To explore the relationship between connectivity changes and spatial proximitly, Pearson's correlation analysis yielded a mean correlation coefficient of $r = 0.00554$ and a mean $p = 0.136$. Similarly, Searman's rank correlation yielded a coefficient of $\rho = 0.00201$ with $p = 0.134$. To account for potential nonlinear relationships, the distance correlation (dCor) value was 0.138. For comparison, the session-averaged mean t-statistic across 100 randomly chosen non-conditioned neurons was -5.44, slightly more negative than the conditioned neuron's value of -4.53. The relationship between the t-statistics of conditioned and non-conditioned neurons across sessions is illustrated in Figure 1a. Notably, the Pearson correlation between the CN and non-CN's t-statistics was strong, with a coefficient of 0.716 and a highly significant p-value of 5.65×10^{-5}.



	Subject ID	Date
Session 1	740369	2025-01-09
Session 2	766719	2025-01-22
Session 3	754303	2025-01-22
Session 4	766719	2025-01-27
Session 5	754303	2025-01-27

(a) t-statistics Over Sessions

(b) Pre- and Post-BCI Neuron Connectivity

Figure 1. Comparison of t-Statistics and Neuron Connectivity Across Sessions

Using a 20-pixel distance threshold to identify CN-targeted photostimulation events, 15 of 25 (60%) experimental sessions contained at least one CN-targeted photostim. Among these 15 sessions, CN-targeted photostims comprised between 0.999% and 3.03% of total photostims, with a mean of 1.89%. Across 14 sessions, the mean ratio of pre-BCI to post-BCI CN-targeted photostims was 1.04, with one session recording 8 pre-BCI and 0 post-BCI photostims. Across the 15 sessions, the mean decrease in CN connectivity post-BCI was 26.2%, with a mean difference of 94 neurons. The changes in CN connectivity pre- and post-BCI for the first 5 sessions are shown in Figure 1b.

3.2 Machine Learning Analysis

3.2.1 Pre- vs. Post-BCI Connectivity Classification

From the 25 experimental sessions, we extracted 50 connectivity matrices, with one pre-BCI and one post-BCI matrix per session. Using the four statistical features derived from each matrix as input to our classifier (see Section 2.3.1) and setting a fixed random seed, the classifier achieved an accuracy of 0.54 ± 0.15. To assess variability from random initialization, we repeated the experiment using 10 different random seeds,

resulting in an average accuracy of 0.60 ± 0.136. A fixed seed was used in all subsequent analyses for reproducibility. Applying PCA to the full connectivity matrices and retaining the top 15 principal components, which explained 96% of the total variance, the classifier achieved an accuracy of 0.60 ± 0.11. A comprehensive summary of additional PCA experiments is provided in Table 1. To assess potential class separability in the PCA-reduced space, we visualized the projection onto the first two principal components, as shown in Figure 2.

# of PCs	Explained Variance	Accuracy
10	0.89	0.520 ± 0.098
15	0.96	0.600 ± 0.110
20	0.98	0.580 ± 0.160

Table 1. PCA Performance on Full Connectivity

# of PCs	Explained Variance	Accuracy
10	0.88	0.500 ± 0.063
30	0.99	0.520 ± 0.117
50	1.00	0.700 ± 0.167

Table 2. PCA Performance on CN Connectivity

When the classifier was trained on seven statistical features extracted from the CN connectivity vectors, we achieved an accuracy of 0.52 ± 0.04. PCA with 50 principal components, which explained 100% of the total variance, achieved an accuracy of 0.70 ± 0.167. A summary of additional PCA experiments is provided in Table 2. Similarly, to examine potential class separation, we projected the data onto the first two principal components. The resulting visualization is shown in Figure 3.

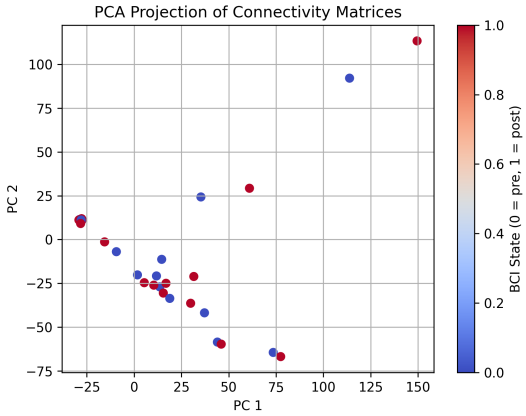


Figure 2. PCA Projection of Full Connectivity

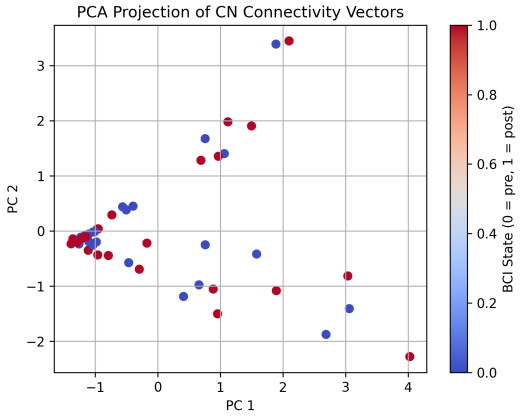


Figure 3. PCA Projection of CN Connectivity

3.2.2 Real vs. Synthetic CN Connectivity Change Classification

From the 25 experimental sessions, we extracted 50 connectivity-change samples, evenly balanced between real and synthetic pairs. The count threshold for feature extraction, as defined in Methods (Section 2.3.2), was optimized via grid search from -1 to 1, which yielded -0.4. Then, using 5-fold cross-validation, the model achieved an average accuracy of 0.90 ± 0.089 in classification of real and synthetic samples.

4 Discussion

4.1 Implications

4.1.1 Statistical Analysis

The significant decrease in CN connectivity from pre- to post-BCI, evidenced by the large negative t-statistic and small p-value, indicates a significant change in neural dynamics following BCI intervention. The change is further supported by the significant decrease in the number of neurons connected to the CN post-BCI, which we determined with a completely different approach. Similarly, non-conditioned neurons exhibited comparable decreases in connectivity, with consistent trends across individual sessions. The strong and statistically significant positive Pearson correlation between CN and non-CN t-statistics suggests that the neuronal connectivity changes are network-wide and not restricted to the conditioned neuron. In addition, the absence of significant linear, monotonic, or nonlinear correlations between connectivity changes and spatial proximity suggests that distance does not influence the observed connectivity changes. This suggests that BCI-induced connectivity changes are governed by global network dynamics rather than local spatial organization. Furthermore, the balanced pre- and post-BCI photostimulation ratio suggests that the observed effects are attributable to biological changes and underlying neural dynamics rather than artifacts such as sampling bias or insufficient data.

4.1.2 Machine Learning Analysis

Initial classification using the four extracted features from the full connectivity matrices and a fixed random seed yielded accuracies that were only slightly above chance and included a large standard deviation, indicating unreliable predictive abilities. Repeating the experiments with different random seeds slightly improved the average accuracy, suggesting the model's moderate sensitivity to initialization. Applying PCA to the full connectivity matrices achieved a moderate improvement over the statistical features, along with reduced variation. However, there was no clear separation between pre- and post-BCI conditions in the PCA visualization, suggesting that the distinctions are very subtle or dimensionality reduction alone may not be sufficient for the task. Further experiments using the CN connectivity vectors and statistical features yielded accuracies near chance level with low variability. In contrast, applying PCA with 50 components to the CN connectivity vectors led to a significant improvement in classification performance. This suggests that PCA is more effective than statistical features as a form of dimensionality reduction and that CN connectivity may contain more distinguishable patterns than full connectivity matrices. The latter is further supported by the increased separation in the PCA visualization. In contrast to the subtle differences observed in the pre/post-BCI classification, the random forest classifier achieved high accuracies in distinguishing between real and synthetic connectivity-change samples using statistical features. The high accuracy indicates the presence of model-detectable structured patterns within the pre- and post-BCI CN connectivity within an experimental session.

4.2 Comparisons

Prior research has reported both increases and decreases in neural connectivity following BCI learning [11, 12]. Our results indicate a global decrease in connectivity post-BCI, suggesting BCI-induced neural plasticity, which we determined with multiple experimental approaches. BCI-induced neural plasticity has been reported in previous research, and our observed reduction in connectivity aligns with theoretical models of synaptic pruning and network refinement, in which learning facilitates the selective weakening of certain connections to make neural processes more efficient [13, 14].

4.3 Limitations

Due to computational concerns and time constraints, we applied a distance threshold that limited our CN-connectivity analysis to 15 out of 25 experimental sessions. Using a larger distance threshold might yield different results. In addition, when classifying pre- and post-BCI neural activity, we treated each segment as a single unit, which may have limited the model's ability to capture session-specific patterns. A more effective approach might involve dividing the segments and performing the analysis on the smaller units.

4.4 Future Research

To deepen our understanding of the neural connectivity dynamics underlying BCI learning, future work will adopt a dynamical systems approach to model connectivity changes over time. This framework could more accurately capture the temporal dynamics of neural plasticity and enable the evaluation of candidate learning rules, including backpropagation and Hebbian learning.

5 Conclusion

Our findings by evaluating neural data from the AIND provide new insight into how cortical circuits adapt during BCI learning, revealing widespread and structured reductions in neural connectivity. These changes were not confined to the conditioned neuron but extended across the network, suggesting a form of global neural plasticity. Random forest classifiers could reliably distinguish real from synthetic CN connectivity changes, indicating the presence of structured, session-specific patterns underlying CN connectivity. While classification of pre- vs. post-BCI connectivity remained challenging, PCA-based models demonstrated improved performance over statistical features, and CN connectivity expressed more underlying patterns than full network connectivity. These results signify the subtlety and complexity of BCI learning-related connectivity changes in the brain and highlight the potential of combining statistical and machine learning analysis.

Acknowledgments

I would like to thank Saskia de Vries and the Allen Institute for Neural Dynamics for providing the dataset and compute resources that made this work possible. Their contributions to neuroscience research are both fascinating and profoundly impactful. I am especially grateful to Professor Matt Golub from the University of Washington for his generous guidance, consistent support, and insightful feedback. His mentorship was pivotal to the development and refinement of this work.

References

[1] C. Pandarinath, D. J. O'Shea, J. Collins, R. Jozefowicz, S. D. Stavisky, J. C. Kao, E. M. Trautmann, M. T. Kaufman, S. I. Ryu, L. R. Hochberg, J. M. Henderson, K. V. Shenoy, L. F. Abbott, and D. Sussillo, "Inferring single-trial neural population dynamics using sequential auto-encoders," *bioRxiv*, 2017. [Online]. Available: https://www.biorxiv.org/content/early/2017/06/20/152884

[2] M. Churchland, J. Cunningham, M. Kaufman, J. Foster, P. Nuyujukian, S. Ryu, and K. Shenoy, "Neural population dynamics during reaching," *Nature*, vol. 487, pp. 51–6, 06 2012.

[3] W. Zong, H. A. Obenhaus, E. R. Skytøen, H. Eneqvist, N. L. de Jong, R. Vale, M. R. Jorge, M.-B. Moser, and E. I. Moser, "Large-scale two-photon calcium imaging in

freely moving mice," *Cell*, vol. 185, no. 7, pp. 1240–1256.e30, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0092867422001970

[4] J. P. Cunningham and B. Yu, "Dimensionality reduction for large-scale neural recordings," *Nature Neuroscience*, vol. 17, no. 11, pp. 1500–1509, Nov. 2014. [Online]. Available: https://doi.org/10.1038/nn.3776

[5] Z. Lu, W. Zhang, T. Le, H. Wang, U. Sümbül, E. T. SheaBrown, and L. Mi, "NetFormer: An Interpretable Model for Recovering Dynamical Connectivity in Neuronal Population Dynamics," in *Proceedings of the International Conference on Learning Representations (ICLR)*, Jan. 2025, spotlight presentation. [Online]. Available: https://openreview.net/forum?id=bcTjW5kS4W

[6] A. C. C. Silveira, A. S. L. ao Marcelo Antunes, M. C. P. Athié, B. F. da Silva, J. ao Victor Ribeiro dos Santos, C. Canateli, M. A. Fontoura, A. Pinto, L. R. Pimentel-Silva, S. H. Avansini, and M. de Carvalho, "Between neurons and networks: investigating mesoscale brain connectivity in neurological and psychiatric disorders," *Frontiers in Neuroscience*, vol. 18, p. 1340345, 2024, review article.

[7] Allen Institute for Neural Dynamics, "Credit Assignment During Learning," https://www.allenneuraldynamics.org/projects/credit-assignment-during-learning, 2025, web project description; Optical BCI task in mouse motor cortex; accessed Jun. 8, 2025.

[8] B. A. Richards and T. P. Lillicrap, "Dendritic solutions to the credit assignment problem," *Current Opinion in Neurobiology*, vol. 54, pp. 28–36, Feb. 2019.

[9] M. Pachitariu, C. Stringer, M. Dipoppa, S. Schröder, L. F. Rossi, H. Dalgleish, M. Carandini, and K. D. Harris, "Suite2p: beyond 10,000 neurons with standard two-photon microscopy," *bioRxiv*, 2017. [Online]. Available: https://www.biorxiv.org/content/early/2017/07/20/061507

[10] T. Zhu, "Analysis on the applicability of the random forest," *Journal of Physics: Conference Series*, vol. 1607, p. 012123, 08 2020.

[11] H. Makino, E. J. Hwang, T. L. Hedrick, and T. Komiyama, "Transformation of cortex-wide emergent properties during motor learning," *Neuron*, vol. 94, no. 4, pp. 880–890.e8, 2017.

[12] K. A. Grigoryan, K. Mueller, M. Wagner, D. Masri, K. J. Pine, A. Villringer, and B. Sehm, "Short-term bci intervention enhances functional brain connectivity associated with motor performance in chronic stroke," *NeuroImage: Clinical*, vol. 46, p. 103772, Mar. 2025, open access under CC BY-NC.

[13] J. J. Daly, R. Cheng, J. Rogers, K. Litinas, A. Hwang, R. Dohrmann, and M. Stokes, "Feasibility of a robot-assisted stroke rehabilitation device," *American Journal of Physical Medicine & Rehabilitation*, vol. 88, no. 6, pp. 450–458, Jun. 2009, open access via PMC.

[14] J. Sakai, "Core concept: How synaptic pruning shapes neural wiring during development and, possibly, in disease," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 117, no. 28, pp. 16 096–16 099, Jun. 2020, published under the PNAS license.