Long-term Intracortical Neural activity and Kinematics (LINK): An intracortical neural dataset for chronic brain-machine interfaces, neuroscience, and machine learning

Hisham Temmar[1], **Yixuan Wang**[2], **Nina Gill**[3], **Nicholas B. Mellon**[4], **Chang Liu**[4],
Luis Hernan Cubillos[5], **Rio I. Parsons**[1], **Joseph T. Costello**[4,6], **Matteo Ceradini**[8],
Madison M. Kelberman[1], **Matthew J. Mender**[1,6], **Aren I. Hite**[1], **Dylan M. Wallace**[5]
Samuel R. Nason-Tomaszewski[1,*] **Matthew S. Willsey**[1,6], **Parag G. Patil**[1,6],
Anne W. Draelos[1,7], **Cynthia A. Chestek**[1,5]

[1]Biomedical Eng., [2]Psychology, [3]Neuroscience, [4]Electrical Engineering and
Computer Science, [5]Robotics, [6]Neurosurgery, [7]Computational Medicine
& Bioinformatics, University of Michigan, Ann Arbor, MI, USA.
[8]The Biorobotics Institute, Sant'Anna School of Advanced Studies, Pisa, Italy.

Abstract

Intracortical brain-machine interfaces (iBMIs) have enabled movement and speech in people living with paralysis by using neural data to decode behaviors in real-time. However, intracortical neural recordings exhibit significant instabilities over time, which poses problems for iBMIs, neuroscience, and machine learning. For iBMIs, neural instabilities require frequent decoder recalibration to maintain high performance, a critical bottleneck for real-world translation. Several approaches have been developed to address this issue, and the field has recognized the need for standardized datasets on which to compare them, but no standard dataset exists for evaluation over year-long timescales. In neuroscience, a growing body of research attempts to elucidate the latent computations performed by populations of neurons. Nonstationarity in neural recordings imposes significant challenges to the design of these studies, so a dataset containing recordings over large time spans would improve methods to account for instabilities. In machine learning, continuous domain adaptation of temporal data is an area of active research, and a dataset containing shift distributions on long time scales would be beneficial to researchers. To address these gaps, we present the LINK Dataset (Long-term Intracortical Neural activity and Kinematics), which contains intracortical spiking activity and kinematic data from 312 sessions of a non-human primate performing a dexterous, 2 degree-of-freedom finger movement task, spanning 1,242 days. We also present longitudinal analyses of the dataset's neural spiking activity and its relationship to kinematics, as well as overall decoding performance using linear and neural network models. The LINK dataset and code are freely available to the public through the dataset website (`https://chesteklab.github.io/LINK_dataset/`).

---

*Currently: Wallace H. Coulter Dept. of Biomedical Engineering, Emory University and Georgia Institute of Technology, Atlanta, GA, USA

1 Introduction

Due to their limited ability to independently engage in activities of daily living, people living with paralysis often report feelings of isolation and lack of personal agency [1–3]. Brain-machine interfaces (BMIs) aim to restore this ability by using recordings from intact neural circuits in the brain and 'decoding' user intent to control an end effector. BMIs have enabled humans and nonhuman primates (NHPs) to control computer cursors, robotic arms, individuated fingers on a virtual hand, and their own (paralyzed) arms through electrical stimulation [4–17]. Additionally, BMIs have helped restore speech in people living with ALS using recurrent-neural network decoders [18–20]. Intracortical brain-machine interfaces (iBMIs) in particular have emerged as a highly promising avenue, due to the high specificity of spiking activity they capture.

However, the specificity of iBMI recordings also leads to one of the primary barriers for real-world translation: instability of neural signals over time. While semi-stable over very short timescales [21–25], intracortical recordings exhibit significant non-stationarities on longer ones [24, 26], due to factors including the foreign body response, scarring, neuronal death, and constant micro-motions in the brain [27–30]. Because of these shifts, current iBMI decoders require frequent recalibration on supervised data, which typically involved a disruptive and lengthy training session. Many approaches have been proposed for addressing this problem, typically aiming to limit the time spent recalibrating or eliminate recalibration altogether through continual unsupervised alignment of neural data [31–42]. Many have recognized the need for standardized benchmarks on which to compare such methods, such as the FALCON dataset [43]. While FALCON covers a broad variety of tasks in the BMI space, the amount of data per task is limited and the timespan is limited to months, not years.

Adjacent to iBMIs, a growing body of neuroscience research uses intracortical recordings to study the dynamics of neural populations, often in the context of motor or cognitive behavior. These studies apply dimensionality reduction and latent variable models to uncover how neural populations perform the computations driving behavior [44–48]. Multiple studies have shown that this latent structure remains stable on a population-level over time, despite recording from unstable individual units [49, 50, 45]. This has led to the hypothesis that behaviorally-relevant computations are embedded in stable manifolds, which persist over time despite turnover at the level of single neurons [40, 49, 39]. However, evidence remains sparse over long timescales, and questions remain about the degree to which these manifolds evolve with experience or become harder to observe with array degradation.

In machine learning, research on domain adaptation for timeseries forecasting generalizes the issue observed in neural decoding to any non-stationary signal over time. Typical forecasting models can be limited by the amount of data available for training, and their training is further complicated by when the underlying distributions of input features and output spaces evolve over time. Domain adaptation methods seek to resolve this problem of changing data distributions [51–53]. For timeseries data, domain adaptation is additionally challenging due to the temporal nature of the data, where patterns learned on an early section of data are not useful for prediction on a later portion. Neural decoders capable of adapting to domain changes in neural and/or kinematic data can indeed be more robust to performance degradation across time [54, 38]. Yet continuous domain adaptation methods for neural network models and times series data have not yet been widely applied, in part due to the lack of available data sampled across both short and long timescales.

To address these needs, we present the Long-term Intracortical Neural activity and Kinematics Dataset (LINK), which contains 312 sessions on 303 days, spanning ~3.5 years of a single non-human primate performing a trial-based dexterous finger movement task. The dataset contains two pre-processed neural features for use in neural decoding tasks and population-level analyses of neural dynamics over time. We hope it can serve as a test bed for developing new domain adaptation approaches for timeseries data more generally. In this work, we describe the dataset and present initial analyses on the quality of neural signals, their relationships with behavior, and some iBMI decoding results.

2 Related Work

Several publicly available datasets have supported progress in BMI decoder performance and robustness, as well as neural population dynamics research. The Neural Latents Benchmark (NLB) [55] introduced a dataset with good task variability, including motor and cognitive tasks, but its temporal span is limited, with most tasks containing single sessions. The more recent FALCON

benchmark [43] includes a wide breadth of tasks and subjects, including data from birds, monkeys, and humans performing multiple motor tasks. However, it is limited in its temporal span and has a sparser distribution of days, making it difficult to study long-term stability.

Beyond benchmark datasets, other relevant open-access datasets have been released through scientific publications. These range from single-session data, such as Brochier et al. [56], to longer spans with lower resolution, such as Gallego-Carracedo et al. [57], which includes data from one subject over 540 days, but only 10 recording sessions. Some studies have examined decoder alignment and stability over extended periods, but the data are either not open-access [49, 40] or cover shorter timescales [58, 18]. LINK fills a key gap by providing long-term, open-access, preprocessed neural and kinematic data with high trial counts, suitable for both decoder stability studies and broader neuroscience research.
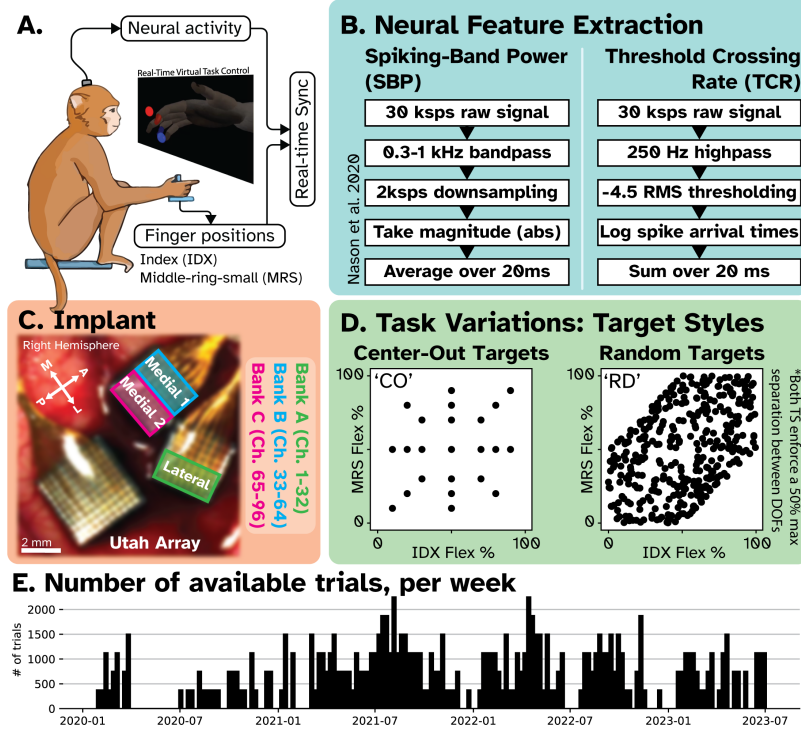
3 The LINK Dataset

Figure 1: **Overview of Experimental Setup. A.** A non-human primate (Monkey N) performed a 2 degree-of-freedom (DOF) finger movement task where he moved his fingers to match targets presented over a virtual hand on a screen. We simultaneously recorded neural activity from 96 channels of a Utah microelectrode array and the positions (using flex sensors) of the two finger groups, index and middle-ring-small. Note that the illustration depicts Monkey N performing the task with his right hand, but in reality he performed the task with his left. **B.** In this dataset, we provide two standard neural features used in iBMI decoding, spiking-band power and threshold crossing rates. The process for extracting each feature (per channel) is shown here. **C.** Monkey N was implanted with 3 Utah arrays, 2 of which were implanted in the motor cortex. Given available hardware, we were able to record 96 channels simultaneously, specified by the colored rectangles. **D.** We included two variations to the task in this dataset, determined by the pattern of targets presented. The first is 'center-out' which enforces stereotyped movements 'out' with every other trial being a return to 'rest' (the 50% flexion and extension dot) and the second is 'random', where targets are randomly presented but with a maximum 50% separation of the fingers. **E.** A histogram showing the number of available trials per week over the 1,242 days contained in the dataset.

The **L**ong-term **I**ntracortical **N**eural and **K**inematics (LINK) Dataset contains data from a single non-human primate, Monkey N, performing a trial-based, 2 degree-of-freedom (DOF) finger task, recorded

on 303 days spanning 1,242 days. The dataset comprises 312 'sessions' of 375 (mostly) contiguous trial acquisitions. Each session contains one of two variations to the task, 'center-out' or 'random', described below. On 9 days, two sessions are available as both tasks were performed. Monkey N was implanted with three microelectrode arrays, two in the right motor cortex and one in the sensory cortex. While Monkey N performed the task, we synchronously measured the spiking activity of neurons from 96 motor array electrodes and the kinematics of the 2DOF task. In each session, we provide neural data preprocessed into two neural features in 20ms bins: spiking-band power (SBP) and threshold crossing rate (TCR). We also provide the position and velocity of the 2DOF at the same resolution. The dataset also includes trial and session-specific information, like the date of acquisition, target pattern used for the session, target positions per trial, and trial timestamps. Each session is included as a separate file conforming to the Neurodata without Borders (NWB) standard, which can be loaded using the pyNWB or MatNWB APIs. The dataset is publicly available on the DANDI archive, `dandiarchive.org/dandiset/001201`, and code for loading the dataset and replicating all analyses in this paper is available at `github.com/chesteklab/LINK_dataset`. Additionally, the dataset and code can be accessed through the landing page at `https://chesteklab.github.io/LINK_dataset/`.

3.1 Experimental Setup

Behavioral Task. A nonhuman primate (NHP, Macaca mulatta), Monkey N, was trained to sit in a chair and perform a trial-based, two DOF dexterous finger movement task, shown in Figure 1A. Monkey N's left hand was positioned in a manipulandum, in which he moved two 'finger groups', the index (IDX) and middle-ring-small (MRS) fingers, to control the positions of matching fingers of a virtual hand present on a screen in front of him. The positions of these finger groups along their respective movement arcs were measured using flex sensors calibrated so that a 0 reading was full extension and 1 was full flexion for each DOF. The trial-based task was as follows: Monkey N was presented with two spherical targets at the beginning of each trial, and had to move his fingers so that the virtual finger groups were in each target (see Figure 1A). After holding each finger group in the targets for 750 ms, the trial was considered a success, and he received a juice reward. If unable to acquire targets after a set time (typically 10 seconds), the trial was considered a failure, and the next set of targets was shown. In this dataset, only successful trials are included. Trials were performed in continuous blocks, called 'sessions'. During a session, targets were presented in different patterns of succession, referred to as 'target styles'. The LINK dataset contains two target styles, *'center-out'* (CO), which mimics the center-out-and-back pattern described in previous studies [59, 60], and *'random'* (RD), in which targets were pseudo-randomly chosen. Both are shown in Figure 1D. In both cases, the range of movement for split targets (e.g., IDX flex, MRS extended) was limited to 50% of the full movement arc, as Monkey N could not reliably extend past this due to natural dexterity. Please see Appendix A for more details on the behavioral task.

Hardware and Feature Extraction. Monkey N was implanted with 3 Utah microelectrode arrays (Blackrock Neurotech, Salt Lake City, UT, USA). One 10×10 array was implanted in the somatosensory cortex, and two 8×8 arrays were implanted in the hand area of the motor cortex based on anatomical targets, shown in Figure 1C and previously described in [61, 62]. The protocols in this study were approved by the Institutional Animal Care and Use Committee at the University of Michigan. Due to hardware limitations, the dataset contains a subset of 96 out of the 128 available motor cortex channels and no somatosensory channels. These were recorded in 32-channel banks, indicated in Figure 1C, labeled by their relative laterality along the motor cortex. For further details on the implants, refer to Appendix A. Neural activity was recorded at 30 kHz by a Cerebus Neural Signal Processor (Blackrock Neurotech), which then extracted two neural features per channel: spiking-band power (SBP) and threshold crossings (TC), outlined in Figure 1B and described in Appendix A. Spiking-band power applies a 300-1000 Hz bandpass filter to each channel, a band which contains the majority of spiking activity, as shown in [63–65]. Threshold crossings are measured by recording time points at which a channel's voltage passes a $-4.5 \times RMS$ (root mean square) threshold set per-channel per day. Both of these features are then binned (averaging for SBP, summing for TCs) into 20ms bins. These binned features will be referred to as SBP and threshold crossing rate (TCR). The recordings included in the dataset begin on day 349 post-implant, up to 1591 days post-implant.

3.2 Data Selection

Over the ~3.5 years that Monkey N's arrays were active and usable for real-time decoding, he typically performed the 2-DOF task 5 days a week (excluding holidays and breaks). However, many were unsuitable for this dataset due to factors like task variation, modification to recording paradigms, high noise, or poor NHP behavior (e.g., low motivation that day). To limit external variability, we performed a three-stage dataset curation process: First, we reviewed all experimental notes for Monkey N and identified candidate blocks of contiguous trials (sessions) for each day, if any were present. Second, we loaded every candidate session and filtered out unsuccessful trials, trials with the wrong target style, closed-loop trials, and trials whose hold times were not 750 ms. To prevent bias in decoder training due to uneven numbers of trials across days, we capped the number of trials to 375 per available target style per day. At this point, we also extracted and binned the neural, kinematic, and trial data. Finally, we conducted a full manual review of each session, inspecting neural data for artifacts, baseline shifts, and other issues, excluding sessions with unsuitable data. These stages were repeated as needed until we settled on a final version of the dataset. For details on each stage of the process, please see Appendix B.

3.3 Data Format and Contents

Each session is contained in an HDF5 file compliant with the Neurodata Without Borders (NWB, [66]) standard. NWB aims to define a common data standard to allow for collective development of tools for neurophysiological data. In each file, we provide the timeseries data (SBP, TCR, behavior, and experiment time) and relevant trial information (target positions, trial start times, lengths, target style). We also provide mappings of the channels to their locations on each array, along with approximate impedance measurements per channel when available. A visual diagram of the channel mappings to electrode positions is available in Figure S1. For ease of use, we have also included code in the dataset repository that converts the NWB files into simple Python dictionaries. Data and associated metadata were organized into a BIDS-like ([67]) structure and uploaded to the DANDI archive ([RRID:SCR_017571], `https://dandiarchive.org/dandiset/001201`), using the Python command line tool `dandi-cli`. Further details on data format and contents are included in Appendix B.

4 Analyses

Here we present several analyses that highlight potential applications of the LINK dataset. Please refer to Appendix C for detailed descriptions of the methods used in these analyses.

4.1 Characterizing Neural Activity Over Time

To our knowledge, this is the largest and time-dense set of neural recordings from a Utah array ever released. As such, we aimed to describe how the neural data evolved over the lifespan of the array, irrespective of behavior. We started by measuring the average SBP of each channel, for each session, shown in Figure 2A. Average SBP across channels decreased gradually by approximately 0.033 μV per month, as estimated by a linear model. There are periods where groups of channels exhibit irregular shifts (e.g. ~2022-07 to ~2022-11), but these did not affect overall decoding performance. During iBMI experiments, we can roughly estimate the number of active channels by counting those with mean TCR > 1 Hz, shown in Figure 2B. Note that this appears to decrease over time. We then used the SBP of these active channels to calculate participation ratio (PR) per day, shown in Figure 2C. Participation ratio roughly corresponds to the number of dimensions required to explain 80% of the total variance in the neural population, as described in [68]. Generally, 10 dimensions were sufficient to capture the majority of the variance in active channels, and a linear model estimated that PR decreased by ~0.07 PR per month. In Figures S4 and S5, we grouped average SBP of each channel by electrode location (32 channel banks) and by active/inactive channels. We observed lower overall activity on the 'lateral' channels and lower SBP in 'inactive' channels over the course of the dataset. All categories saw steady decreases in average SBP over time.

Additionally, we investigated the evolution of neural activity at a population level. We first concatenated the SBP across all sessions, z-scored per channel, and fit a PCA transform across the entire dataset. Then, we measured the per-day centroid of the top 3 PCs and plotted them in Figure 2D.

We also measured the centroids when grouping the data per-quarter (of a year), shown in larger points, and plotted the standard deviation of these larger groupings as shaded ellipsoids. Visually, these centroids appear to move throughout the PCA space over time, and except for two quarters, the standard deviation of the top 3 PCs does not exhibit large changes.

We then examined the behavior of the neural population during different movements. To do this, we fit PCA transforms to per-channel z-scored neural data within each day, and grouped trials by year and required movement direction (as shown in Figure 2E). Trajectories through the PC space were then averaged over time-aligned trials (by max jerk pre-movement onset) within each year and movement direction. This was performed over both center-out and random trials. The average trajectories for 2020 and 2023 (up to the end of the dataset, containing ~6 months) are shown in Figure 2E. The trajectories shown remain visually separable at the beginning and end of the dataset. For further details about how we examined neural activity over time, please refer to Appendix C.3.
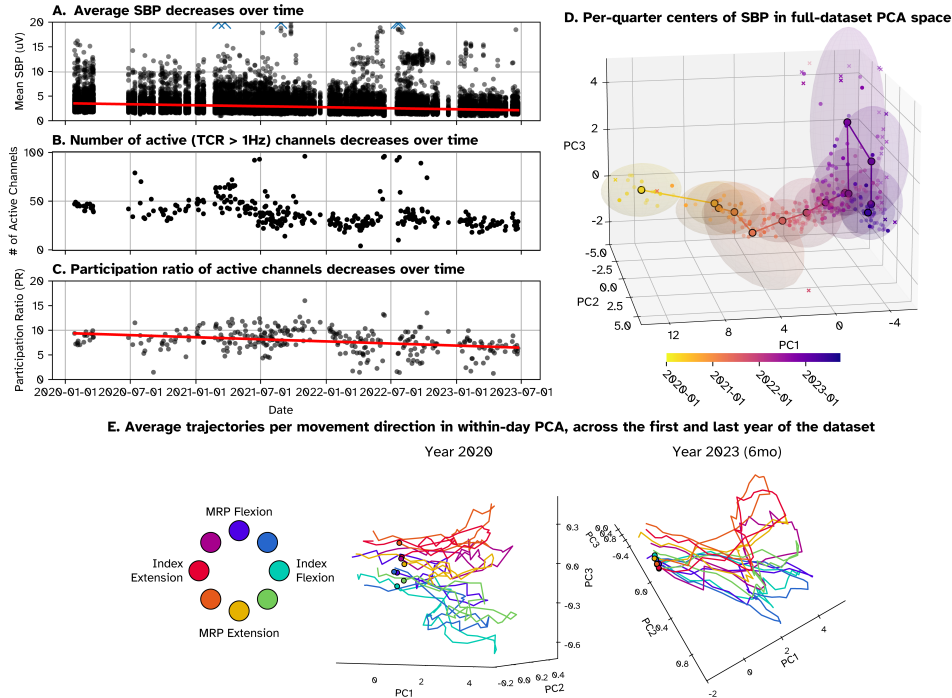

Figure 2: **Changes in neural signals over time.** Here we investigated the change in neural activity over time using various metrics. Additionally, we briefly examine how low-dimensional representations of population-level neural activity relates to behavior of the course of the dataset. **A.** In this plot, each point represents the average SBP of a single channel on a single day. In red, a linear model shows the change in average SBP across channels over time (slope 0.033 μV per month). For ease of visualization, the y-axis is limited, but outliers are marked with blue X's. **B.** We counted the number of 'active' channels each day with an average TCR of >1 Hz. This plot shows the number of active channels on each day included in the dataset. **C.** We measured the participation ratio using the SBP of all 'active' channels for each day and plot them over time. A linear model was fit to the data, shown in red. **D.** We z-scored SBP per-channel and fit a PCA transformation across the entire dataset. The mean of the top 3 PCs per day (small dots) are shown as the small dots, colored by date as shown in the colorbar. The mean of the top 3 PCs per-quarter (of a year) are shown as the larger dots, also colored by date. The shaded ellipsoids represent the standard deviation of the top 3 PC scores grouped by quarter. For visual clarity, axis limits were restricted to what is shown, but any outliers are represented with X's at the axis limits, colored by date. **E.** On each day, SBP was z-scored per-channel and a PCA transform was applied. Trials were then grouped by general direction of required movement and aligned by max jerk before movement onset. The two plots show the average trajectories in the top 3 PCs for each movement direction in the first and last year.

4.2 Preferred Tuning Direction

One approach to investigating how the relationship between neural activity and behavior changes over time is via the individual relationship between a channel and each degree of freedom. We calculated the *preferred tuning vector* of each channel for each session by fitting independent linear regressions between each channel and the position of both degrees of freedom, and taking the weights of each as a vector, as described in Appendix C.2. These vectors are then decomposed into a 'tuning direction' and 'tuning strength', as shown in Figure 3A. Figure 3B visualizes the preferred tunings of two channels over the entire dataset, one point for each day colored by date. Both channels were considered active, with average TCRs >1 Hz across the dataset. Channel 7 was strongly tuned to MRS flexion throughout the dataset, whereas channel 32 exhibited large shifts in tuning angle. The preferred tuning angles and strengths of all channels on all days are shown in Figures 3C and D, respectively. Per-channel tuning angles and strengths generally appear consistent throughout the dataset; however, many channels have very low tuning strengths throughout the lifespan of the arrays. Such channels also typically showed low average TCRs per day, as shown in Figure S3. To quantify the range of tuning angles per channel across time, we measured the circular median and circular IQR (inter-quartile range) of tuning angle per channel over the whole dataset and the median and IQR of tuning strength per channel over the whole dataset (Figure 3E, described in Appendix C.2). From these plots, it appears that channels in the medial array (see Figure 1) are mainly tuned to MRS flexion, with some IDX flexion, while channels in the lateral array are more tuned to extension of both DOF. In most channels, the circular IQR of tuning angles was small, suggesting their tunings did not change much throughout the dataset.

4.3 Neural Decoding

To validate each session and corresponding decoders, each session was split into 300 trials of training data and 75 hold-out trials. The SBP of the training data was z-scored and used to train a ridge regression (RR) and a long short-term memory network (LSTM, [69]). More details on decoder training can be found in Appendix C.1. These decoders predicted both position and velocity of the two finger groups, although in most BMI applications, only velocity is used for closed-loop control. Prediction accuracy (R^2) across DOFs for each day can be seen in Figure 4A. Within-day decoding accuracy remains relatively stable across time, and the majority of shifts in accuracy are reflected in both decoders, suggesting these shifts may be due to changes in modulation strength highlighted in Figure 3B. Prediction accuracy by DOF (and position and velocity) is shown in Figure 4B. In general, the position accuracies of the two finger groups outperformed velocities, and MRS predictions outperformed IDX predictions, suggesting there may be more information regarding the MRS finger group in the neural data than IDX.

To visualize the problem of decoder failure over time, all decoders were evaluated on all holdouts within 100 days of training. All decoder predictions were grouped by the number of days out from decoder training, starting with day 0 (decoder training). Figure 4D shows average accuracy across these groups. Decoding accuracy across DOFs was averaged over all predictions. Each performance was labeled with the dates of decoder training, hold-out sets, and the number of days from training. LSTM outperformed RR on all days, but both followed a similar decay curve.

Many BCI decoding algorithms use multiple days of pre-training to improve performance. To test this, we trained LSTMs on up to 5 consecutive session and evaluated their performance on all sessions within 180 days of day 0 (latest day). We excluded days with multiple sessions and the first 10 sessions of the dataset. Note that since sessions are not evenly distributed the time span covered by training sessions varies. Performances (R^2) were grouped by relative day of prediction (day k from training) and by number of sessions included in training (1-5). These are plotted in Figure 5A. Multi-day training, reduced the initial decrease in performance (within ~5 days of training), improving performance over time, but did not seem to impact the longer term decay, and slightly decreases day 0 performance. A double-exponential fit (Fig. 5A) to the single-session and 5-session conditions over all evaluation sessions (within 180 days of day 0) show steeper decays in the single-session case, confirming our observations: the initial decay decreases with more training sessions, but the long-term decay is not significantly affected.

A complementary approach to multi-session pretraining is continual fine-tuning small amounts of (labelled) incoming data, gradually adapting the decoder to shifting neural signals while removing the need to collect entire training sessions every day. To simulate this, we trained an LSTM on a seed day,
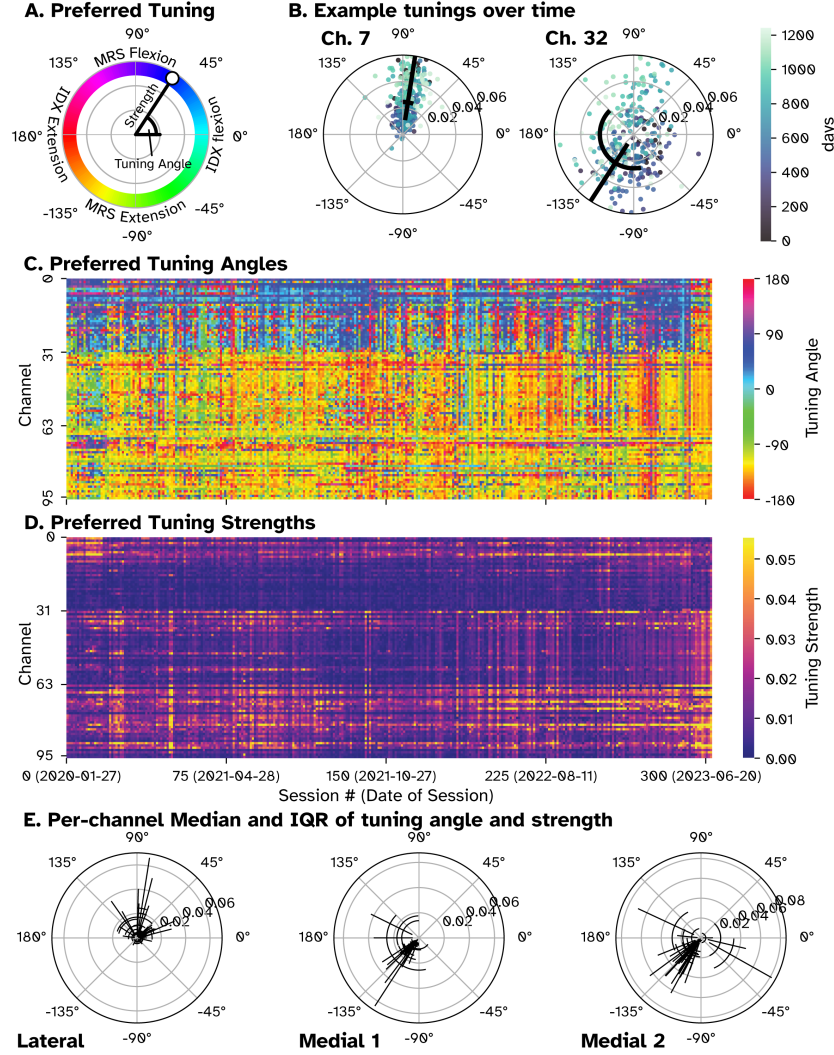
7

Figure 3: **Single-channel preferred tuning over time.** We investigated the relationship between the SBP of individual neural channels to behavior over the entire dataset by measuring the preferred tuning vectors of each channel for every session. In general, channels tended to stay tuned to the same direction throughout the dataset, with varying tuning strengths. **A.** Toy example illustrating how to interpret preferred tuning for a channel. Preferred tuning is a vector, measured on SBP, whose angle denotes the 'preferred' movement direction, and whose magnitude indicates the 'strength' of tuning towards this angle. **B.** Here we show the preferred tuning of two channels throughout the dataset. Each point shows the tuning of a single day, colored to indicate its temporal position in the dataset. The black crosses are centered at the median and circular median of the tuning strength and angle, respectively, while the size of the line and arc indicate the IQR (inter-quartile range) and circular IQR of the tuning strength and angle, respectively. **C.** Preferred tuning angles for all channels on all days. The preferred angle for each channel on each day in the dataset is colored according to the cyclical color bar and presented as a heatmap **D.** Preferred tuning strengths for all channels on all days. The preferred tuning strength for each channel on each day is colored according to the colorbar and presented as a heatmap. Any breaks between sessions are removed for visual clarity **E.** Per-channel tuning ranges across the dataset, grouped by bank. Channels were recorded in 32-channel banks, one from one array, and two from another. Here, we capture the range of tunings measured for each channel by plotting the circular median/median of the preferred angle/strength for each channel, and the circular IQR/IQR of each channel as error bars. We separated each bank into separate plots.
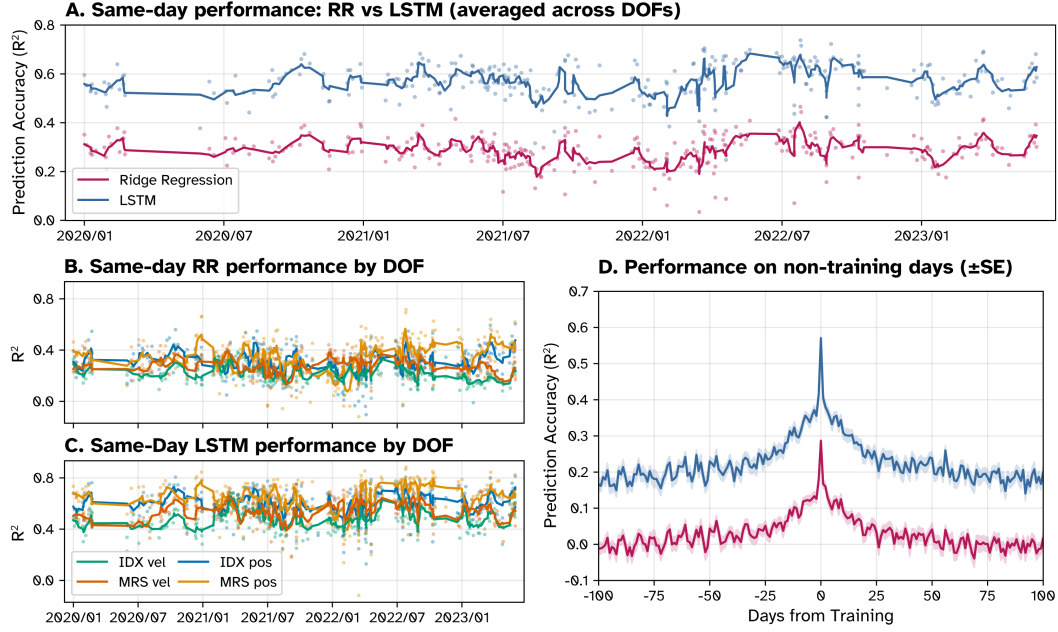
8

Figure 4: **Baseline decoding performance over the whole dataset.** Each session was split into a 300-trial training dataset and 75-trial hold-out, used to train two decoders, ridge regression (RR) and a long-short term RNN (LSTM). Prediction accuracy was measured using R^2. **A.** R^2 of each decoder tested on the same-day hold-out set. Points show per day R^2 averaged across DOF and the line shows the 5-day rolling average. **B-C.** Same-day performance separated by DOF for RR (B) and LSTM (C). **D.** Decoders were evaluated on hold-out sets within 100 days of training. Each line shows the average R^2 across all evaluations K days from decoder training, separated by decoder. Day 0 R^2 for each decoder is the average of the points in A. $\pm$ Standard errors are shown as shaded regions. Neither RR nor LSTM generalized well over time.

and then, for up to 200 days following training, we fine-tuned the LSTM on each consecutive session using 30 s to 5 minutes of training data. We repeated this for 20 seed days approximately evenly spaced across the dataset. We then plotted the average R^2 across the 20 seed days for each relative day from initial training, as shown in 5B. Fine-tuning with as little as 30 s of data helped reduce both the short and long-term decays in accuracy. Note that since fine-tuning was performed continually (e.g. a decoder fine-tuned on day 200 had also been fed data from all prior available sessions), performance generally improved gradually over time. In fact, fine-tuning on 120s or 300s led to later decoders outperforming the baseline day 0 decode. These results suggest that daily finetuning with relatively small amounts of labeled data can help with recovering performance.

5 Limitations, Discussion, and Conclusions

To our knowledge, LINK is one of the largest and longest spanning public releases of behaviorally-relevant intracortical data to date. The dataset contains two standard neural features (for iBMI decoding and neural population analysis), with two different tasks and details about each of the 96 channels. This data is provided for hundreds of trials per day, on hundreds of days, spanning almost 4 years. We believe the LINK dataset fills a critical gap and will enable new and interesting research in BMI decoding and neuroscience. However, the dataset has multiple limitations. First, the dataset only contains one subject, which may bias data. Secondly, the dataset contains multiple extended time gaps, some of which span months, mostly due to clusters of days failing to meet quality criteria (see Supplementary Materials). Additionally, due to storage limitations, raw voltage signals were not recorded for the majority of the sessions included in the dataset, limiting its use for developing new neural features or decoding methods which use raw data. Finally, due to the trial-based, constrained nature of the task, the actual task itself is not as behaviorally rich nor broadly labeled as might be desired for the development of foundational models for neuroscience [70, 71].

A. Effects of multi-session LSTM training

- - - - 1 session double-exp fit: $0.229e^{-0.183x} + 0.268e^{-0.006x}$
- - - - 5 session double-exp fit: $0.134e^{-0.149x} + 0.374e^{-0.004x}$

# of training sets
— 1 session
— 2 sessions
— 3 sessions
— 4 sessions
— 5 sessions

R^2

Relative days since day 0 (day K)

B. Effects of daily fine-tuning on limited data

R^2

Fine-tuning data size
— 30s — 120s
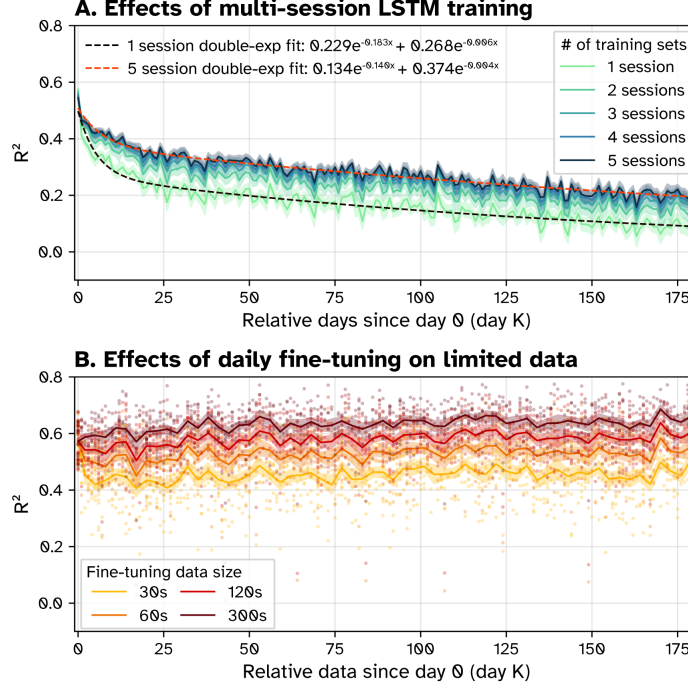— 60s — 300s

Relative data since day 0 (day K)

Figure 5: **Modified training paradigms for recalibration.** Here we test the ability of two training paradigms to improve the performance of LSTM decoders over time. **A.** On 283 days, 5 LSTM decoders were trained, using from 1-5 past sessions of data (current day inclusive). The plot shows the average prediction R^2 of the decoders tested on future sessions, across all models, grouped by the number of sessions included in the training data. **B.** We also tested performance when the LSTM model was fine-tuned on small amounts of future data every day, up to 200 days.

Figure 2 demonstrated that across all channels, neural activity generally decreased over time. This has been previously observed in [21], and is expected, as the array is gradually encapsulated by scar tissue over time [27]. However, this did not correspond to decreases in decoding performance throughout the dataset, suggesting that sufficient behavioral information remained for high-accuracy decoding. This is further supported by observations where, when projected into a low-dimensional neural space, neural trajectories for different directions remained separable throughout. More advanced and neuro-focused methods for latent feature extraction/dimensionality reduction (e.g. LFADS [48, 39], jPCA [47], dPCA [46], Isomap [72, 73]) may provide further insight into the evolution of population-level activity over time. In future datasets, it would be interesting to see if the gradual decrease in neural activity has an earlier observable impact on decoding accuracy for more complex tasks.

In Figure 3, we observed that many channels maintained a relatively stable preferred tuning direction over time (while tuning strength varied). Given that SBP and TCR capture the spiking activity of individual neurons close to the electrode [63], this stability in tuning direction could suggest that many electrodes are recording from the same neurons. We also observed that channels located on the same array tended to exhibit similar tuning angles, as could be expected from a somatotopic organization [74]. A potentially novel observation from the preferred tuning directions is the apparent separation of MRS flexion and extension tunings by array, but this data comes from only a single subject and task, and such a claim would require extensive investigation across several subjects.

In general, single-day decoding performance was fairly consistent throughout the lifespan of the array. While the LSTM outperformed RR on every day, we noted that there were baseline shifts in decoding performance that impacted both decoders, suggesting there are underlying shifts in the behaviorally-relevant information available in the neural data. When testing decoders on past and future data, we noted that both RR and LSTM exhibited sharp decreases in performance over the short-term, agreeing with previous observations [32, 26]. However, this decreases stabilizes on longer timescales, suggesting that there may be a stable subspace of neural activity that is at least partially behaviorally-relevant and nonlinearly related to behavior (hence the increased performance of the

LSTM decoder). This may also be supported by the shape of the decay being agnostic to the decoder used and the relatively consistent tuning directions of single channels.

Acknowledgments and Disclosure of Funding

References

[1] Wang Dan, Sheilla M Trajera, and Gregory S Ching. Living with amyotrophic lateral sclerosis (ALS): Navigating the various emotional and psychosocial challenges. 13(3):7006–7013.

[2] Stephanie R Cimino, Lauren Cadel, Sara JT Guilcher, Marina Wasilewski, and Sander L Hitzig. Social disconnectedness and perceived social isolation in persons with spinal cord injury/dysfunction living in the community: A scoping review. 46(3):367–389. Publisher: Taylor & Francis.

[3] Maggi A Budd, David R Gater Jr, and Isabella Channell. Psychosocial consequences of spinal cord injury: a narrative review. 12(7):1178. Publisher: MDPI.

[4] Leigh R Hochberg, Mijail D Serruya, Gerhard M Friehs, Jon A Mukand, Maryam Saleh, Abraham H Caplan, Almut Branner, David Chen, Richard D Penn, and John P Donoghue. Neuronal ensemble control of prosthetic devices by a human with tetraplegia. 442(7099):164–171. Publisher: Nature Publishing Group UK London.

[5] Brian M Dekleva, Jeffrey M Weiss, Michael L Boninger, and Jennifer L Collinger. Generalizable cursor click decoding using grasp-related neural transients. 18(4):0460e9. Publisher: IOP Publishing.

[6] Mijail D Serruya, Nicholas G Hatsopoulos, Liam Paninski, Matthew R Fellows, and John P Donoghue. Instant neural control of a movement signal. 416(6877):141–142.

[7] Sam Musallam, BD Corneil, Bradley Greger, Hans Scherberger, and Richard A Andersen. Cognitive control signals for neural prosthetics. 305(5681):258–262. Publisher: American Association for the Advancement of Science.

[8] Leigh R Hochberg, Daniel Bacher, Beata Jarosiewicz, Nicolas Y Masse, John D Simeral, Joern Vogel, Sami Haddadin, Jie Liu, Sydney S Cash, Patrick Van Der Smagt, and others. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. 485(7398):372–375. Publisher: Nature Publishing Group UK London.

[9] Jennifer L Collinger, Brian Wodlinger, John E Downey, Wei Wang, Elizabeth C Tyler-Kabara, Douglas J Weber, Angus J C McMorland, Meel Velliste, Michael L Boninger, and Andrew B Schwartz. High-performance neuroprosthetic control by an individual with tetraplegia. 381(9866):557–564.

[10] Jose M Carmena, Mikhail A Lebedev, Roy E Crist, Joseph E O'Doherty, David M Santucci, Dragan F Dimitrov, Parag G Patil, Craig S Henriquez, and Miguel A L Nicolelis. Learning to control a brain–machine interface for reaching and grasping by primates. 1(2):e42. Publisher: Public Library of Science.

[11] Abidemi Bolu Ajiboye, Francis R Willett, Daniel R Young, William D Memberg, Brian A Murphy, Jonathan P Miller, Benjamin L Walter, Jennifer A Sweet, Harry A Hoyen, Michael W Keith, and others. Restoration of reaching and grasping in a person with tetraplegia through brain-controlled muscle stimulation: a proof-of-concept demonstration. 389(10081):1821.

[12] Chad E. Bouton, Ammar Shaikhouni, Nicholas V. Annetta, Marcia A. Bockbrader, David A. Friedenberg, Dylan M. Nielson, Gaurav Sharma, Per B. Sederberg, Bradley C. Glenn, W. Jerry Mysiw, Austin G. Morgan, Milind Deogaonkar, and Ali R. Rezai. Restoring cortical control of functional movement in a human with quadriplegia. 533(7602):247–250. Publisher: Nature Publishing Group.

[13] Chet T Moritz, Steve I Perlmutter, and Eberhard E Fetz. Direct control of paralysed muscles by cortical neurons. 456(7222):639–642. Publisher: Nature Publishing Group UK London.

[14] Christian Ethier, Emily R Oby, Matthew J Bauman, and Lee E Miller. Restoration of grasp following paralysis through brain-controlled stimulation of muscles. 485(7398):368–371. Publisher: Nature Publishing Group UK London.

[15] Marco Capogrosso, Tomislav Milekovic, David Borton, Fabien Wagner, Eduardo Martin Moraud, Jean-Baptiste Mignardot, Nicolas Buse, Jerome Gandar, Quentin Barraud, David Xing, and others. A brain–spine interface alleviating gait deficits after spinal cord injury in primates. 539(7628):284–288. Publisher: Nature Publishing Group UK London.

[16] Marion Badi, Sophie Wurth, Ilaria Scarpato, Evgenia Roussinova, Elena Losanno, Andrew Bogaard, Maude Delacombaz, Simon Borgognon, Paul Cvancara, Florian Fallegger, and others. Intrafascicular peripheral nerve stimulation produces fine functional hand movements in primates. 13(617):eabg6463. Publisher: American Association for the Advancement of Science.

[17] Matthew S. Willsey, Nishal P. Shah, Donald T. Avansino, Nick V. Hahn, Ryan M. Jamiolkowski, Foram B. Kamdar, Leigh R. Hochberg, Francis R. Willett, and Jaimie M. Henderson. A high-performance brain–computer interface for finger decoding and quadcopter game control in an individual with paralysis. *Nature Medicine*, 31(1):96–104, January 2025. Publisher: Nature Publishing Group.

[18] Francis R. Willett, Erin M. Kunz, Chaofei Fan, Donald T. Avansino, Guy H. Wilson, Eun Young Choi, Foram Kamdar, Matthew F. Glasser, Leigh R. Hochberg, Shaul Druckmann, Krishna V. Shenoy, and Jaimie M. Henderson. A high-performance speech neuroprosthesis. 620(7976):1031–1036. Number: 7976 Publisher: Nature Publishing Group.

[19] Sean L. Metzger, Kaylo T. Littlejohn, Alexander B. Silva, David A. Moses, Margaret P. Seaton, Ran Wang, Maximilian E. Dougherty, Jessie R. Liu, Peter Wu, Michael A. Berger, Inga Zhuravleva, Adelyn Tu-Chan, Karunesh Ganguly, Gopala K. Anumanchipalli, and Edward F. Chang. A high-performance neuroprosthesis for speech decoding and avatar control. 620(7976):1037–1046. Number: 7976 Publisher: Nature Publishing Group.

[20] Nicholas S. Card, Maitreyee Wairagkar, Carrina Iacobacci, Xianda Hou, Tyler Singer-Clark, Francis R. Willett, Erin M. Kunz, Chaofei Fan, Maryam Vahdati Nia, Darrel R. Deo, Aparna Srinivasan, Eun Young Choi, Matthew F. Glasser, Leigh R. Hochberg, Jaimie M. Henderson, Kiarash Shahlaie, Sergey D. Stavisky, and David M. Brandman. An Accurate and Rapidly Calibrating Speech Neuroprosthesis. *New England Journal of Medicine*, 391(7):609–618, August 2024. Publisher: Massachusetts Medical Society _eprint: https://www.nejm.org/doi/pdf/10.1056/NEJMoa2314132.

[21] Cynthia A. Chestek, Vikash Gilja, Paul Nuyujukian, Justin D. Foster, Joline M. Fan, Matthew T. Kaufman, Mark M. Churchland, Zuley Rivera-Alvidrez, John P. Cunningham, Stephen I. Ryu, and Krishna V. Shenoy. Long-term stability of neural prosthetic control signals from silicon cortical arrays in rhesus macaque motor cortex. 8(4):045005.

[22] J D Simeral, S-P Kim, M J Black, J P Donoghue, and L R Hochberg. Neural control of cursor trajectory and click by a human with tetraplegia 1000 days after implant of an intracortical microelectrode array. 8(2):025027.

[23] Mijail Serruya, Nicholas Hatsopoulos, Matthew Fellows, Liam Paninski, and John Donoghue. Robustness of neuroprosthetic decoding algorithms. 88:219–228. Publisher: Springer.

[24] János A Perge, Mark L Homer, Wasim Q Malik, Sydney Cash, Emad Eskandar, Gerhard Friehs, John P Donoghue, and Leigh R Hochberg. Intra-day signal instabilities affect decoding performance in an intracortical neural interface system. 10(3):036004. Publisher: IOP Publishing.

[25] Cynthia A Chestek, Aaron P Batista, Gopal Santhanam, Byron M Yu, Afsheen Afshar, John P Cunningham, Vikash Gilja, Stephen I Ryu, Mark M Churchland, and Krishna V Shenoy. Single-neuron stability during repeated reaching in macaque premotor cortex. 27(40):10742–10750.

[26] John E. Downey, Nathaniel Schwed, Steven M. Chase, Andrew B. Schwartz, and Jennifer L. Collinger. Intracortical recording stability in human brain-computer interface users. *Journal of neural engineering*, 15(4):046016, August 2018. Place: England.

[27] Ashley N Dalrymple, Sonny T Jones, James B Fallon, Robert K Shepherd, and Douglas J Weber. Overcoming failure: improving acceptance and success of implanted neural interfaces. 11(1):6. Publisher: Springer.

[28] Sajishnu P Savya, Fan Li, Stephanie Lam, Steven M Wellman, Kevin C Stieger, Keying Chen, James R Eles, and Takashi DY Kozai. In vivo spatiotemporal dynamics of astrocyte reactivity following neural electrode implantation. 289:121784. Publisher: Elsevier.

[29] Alejandro Carnicer-Lombarte, Shao-Tuan Chen, George G Malliaras, and Damiano G Barone. Foreign body reaction to implanted biomaterials and its impact in nerve neuroprosthetics. 9:622524. Publisher: Frontiers Media SA.

[30] Scott F Lempka, Matthew D Johnson, Michael A Moffitt, Kevin J Otto, Daryl R Kipke, and Cameron C McIntyre. Theoretical analysis of intracortical microelectrode recordings. 8(4):045006. Publisher: IOP Publishing.

[31] P Zhang, W Li, X Ma, J He, J Huang, and Q Li. Feature-selection-based transfer learning for intracortical brain–machine interface decoding. 29:60–73.

[32] David Sussillo, Sergey D. Stavisky, Jonathan C. Kao, Stephen I. Ryu, and Krishna V. Shenoy. Making brain–machine interfaces robust to future neural variability. 7(1):13749. Publisher: Nature Publishing Group.

[33] Thomas Hosman, Tsam Kiu Pun, Anastasia Kapitonava, John D. Simeral, and Leigh R. Hochberg. Months-long high-performance fixed LSTM decoder for cursor control in human intracortical brain-computer interfaces. In *2023 11th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 1–5. ISSN: 1948-3554.

[34] Chaofei Fan, Nick Hahn, Foram Kamdar, Donald Avansino, Guy Wilson, Leigh Hochberg, Krishna V Shenoy, Jaimie Henderson, and Francis Willett. Plug-and-play stability for intracortical brain-computer interfaces: A one-year demonstration of seamless brain-to-text communication. In A. Oh, T. Naumann, A. Globerson, K. Saenko, M. Hardt, and S. Levine, editors, *Advances in Neural Information Processing Systems*, volume 36, pages 42258–42270. Curran Associates, Inc.

[35] Yu Qi, Bin Liu, Yueming Wang, and Gang Pan. Dynamic ensemble modeling approach to nonstationary neural decoding in brain-computer interfaces. In *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.

[36] Shoeb Shaikh, Rosa So, Tafadzwa Sibindi, Camilo Libedinsky, and Arindam Basu. Sparse ensemble machine learning to improve robustness of long-term decoding in iBMIs. 28(2):380–389.

[37] Yuanrui Dong, Dingyin Hu, Shirong Wang, and Jiping He. Decoder calibration framework for intracortical brain-computer interface system via domain adaptation. 81:104453.

[38] Guy H. Wilson, Francis R. Willett, Elias A. Stein, Foram Kamdar, Donald T. Avansino, Leigh R. Hochberg, Krishna V. Shenoy, Shaul Druckmann, and Jaimie M. Henderson. Long-term unsupervised recalibration of cursor BCIs. Pages: 2023.02.03.527022 Section: New Results.

[39] Brianna M. Karpowicz, Yahia H. Ali, Lahiru N. Wimalasena, Andrew R. Sedler, Mohammad Reza Keshtkaran, Kevin Bodkin, Xuan Ma, Lee E. Miller, and Chethan Pandarinath. Stabilizing brain-computer interfaces through alignment of latent dynamics, November 2022. Pages: 2022.04.06.487388 Section: New Results.

[40] Alan D. Degenhart, William E. Bishop, Emily R. Oby, Elizabeth C. Tyler-Kabara, Steven M. Chase, Aaron P. Batista, and Byron M. Yu. Stabilization of a brain–computer interface via the alignment of low-dimensional spaces of neural activity. *Nature Biomedical Engineering*, 4(7):672–685, July 2020. Number: 7 Publisher: Nature Publishing Group.

[41] Ali Farshchian, Juan A. Gallego, Joseph P. Cohen, Yoshua Bengio, Lee E. Miller, and Sara A. Solla. Adversarial domain adaptation for stable brain-machine interfaces.

[42] Xuan Ma, Fabio Rizzoglio, Kevin L Bodkin, Eric Perreault, Lee E Miller, and Ann Kennedy. Using adversarial networks to extend brain computer interface decoding accuracy over time. 12:e84296. Publisher: eLife Sciences Publications, Ltd.

[43] Brianna M. Karpowicz, Joel Ye, Chaofei Fan, Pablo Tostado-Marcos, Fabio Rizzoglio, Clay Washington, Thiago Scodeler, Diogo de Lucena, Samuel R. Nason-Tomaszewski, Matthew J. Mender, Xuan Ma, Ezequiel M. Arneodo, Leigh R. Hochberg, Cynthia A. Chestek, Jaimie M. Henderson, Timothy Q. Gentner, Vikash Gilja, Lee E. Miller, Adam G. Rouse, Robert A. Gaunt, Jennifer L. Collinger, and Chethan Pandarinath. Few-shot algorithms for consistent neural decoding (FALCON) benchmark. 37:76578–76615.

[44] Byron M. Yu, John P. Cunningham, Gopal Santhanam, Stephen I. Ryu, Krishna V. Shenoy, and Maneesh Sahani. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *Journal of Neurophysiology*, 102(1):614–635, July 2009. Publisher: American Physiological Society.

[45] Juan A. Gallego, Matthew G. Perich, Lee E. Miller, and Sara A. Solla. Neural Manifolds for the Control of Movement. *Neuron*, 94(5):978–984, June 2017.

[46] Dmitry Kobak, Wieland Brendel, Christos Constantinidis, Claudia E Feierstein, Adam Kepecs, Zachary F Mainen, Xue-Lian Qi, Ranulfo Romo, Naoshige Uchida, and Christian K Machens. Demixed principal component analysis of neural population data. *eLife*, 5:e10989, April 2016. Publisher: eLife Sciences Publications, Ltd.

[47] Mark M. Churchland, John P. Cunningham, Matthew T. Kaufman, Justin D. Foster, Paul Nuyujukian, Stephen I. Ryu, and Krishna V. Shenoy. Neural population dynamics during reaching. *Nature*, 487(7405):51–56, July 2012. Number: 7405 Publisher: Nature Publishing Group.

[48] Chethan Pandarinath, Daniel J. O'Shea, Jasmine Collins, Rafal Jozefowicz, Sergey D. Stavisky, Jonathan C. Kao, Eric M. Trautmann, Matthew T. Kaufman, Stephen I. Ryu, Leigh R. Hochberg, Jaimie M. Henderson, Krishna V. Shenoy, L. F. Abbott, and David Sussillo. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods*, 15(10):805–815, October 2018. Number: 10 Publisher: Nature Publishing Group.

[49] Juan A. Gallego, Matthew G. Perich, Raeed H. Chowdhury, Sara A. Solla, and Lee E. Miller. Long-term stability of cortical population dynamics underlying consistent behavior. *Nature Neuroscience*, 23(2):260–270, February 2020. Number: 2 Publisher: Nature Publishing Group.

[50] Matthew G. Perich, Sara Conti, Marion Badi, Andrew Bogaard, Beatrice Barra, Kanaka Rajan, Jocelyne Bloch, Gregoire Courtine, Silvestro Micera, Marco Capogrosso, and Tomislav Milekovic. Motor cortical dynamics are shaped by multiple distinct subspaces during naturalistic behavior, July 2024. Pages: 2020.07.30.228767 Section: New Results.

[51] Abolfazl Farahani, Sahar Voghoei, Khaled Rasheed, and Hamid R Arabnia. A brief review of domain adaptation. pages 877–894. Publisher: Springer.

[52] Xiaoyong Jin, Youngsuk Park, Danielle Maddix, Hao Wang, and Yuyang Wang. Domain adaptation for time series forecasting via attention sharing. In *International Conference on Machine Learning*, pages 10280–10297. PMLR.

[53] Zehao Xiao and Cees GM Snoek. Beyond model adaptation at test time: A survey.

[54] Eva L Dyer, Mohammad Gheshlaghi Azar, Matthew G Perich, Hugo L Fernandes, Stephanie Naufel, Lee E Miller, and Konrad P Körding. A cryptography-based approach for movement decoding. 1(12):967–976. Publisher: Nature Publishing Group UK London.

[55] Felix Pei, Joel Ye, David Zoltowski, Anqi Wu, Raeed H. Chowdhury, Hansem Sohn, Joseph E. O'Doherty, Krishna V. Shenoy, Matthew T. Kaufman, Mark Churchland, Mehrdad Jazayeri, Lee E. Miller, Jonathan Pillow, Il Memming Park, Eva L. Dyer, and Chethan Pandarinath. Neural Latents Benchmark '21: Evaluating latent variable models of neural population activity, January 2022. arXiv:2109.04463 [cs].

[56] Thomas Brochier, Lyuba Zehl, Yaoyao Hao, Margaux Duret, Julia Sprenger, Michael Denker, Sonja Grün, and Alexa Riehle. Massively parallel recordings in macaque motor cortex during an instructed delayed reach-to-grasp task. *Scientific Data*, 5(1):180055, April 2018. Publisher: Nature Publishing Group.

[57] Cecilia Gallego-Carracedo, Matthew G Perich, Raeed H Chowdhury, Lee E Miller, and Juan Álvaro Gallego. Local field potentials reflect cortical population dynamics in a region-specific and frequency-dependent manner. *eLife*, 11:e73155, August 2022. Publisher: eLife Sciences Publications, Ltd.

[58] Tsam Kiu Pun, Mona Khoshnevis, Tommy Hosman, Guy H. Wilson, Anastasia Kapitonava, Foram Kamdar, Jaimie M. Henderson, John D. Simeral, Carlos E. Vargas-Irwin, Matthew T. Harrison, and Leigh R. Hochberg. Measuring instability in chronic human intracortical neural recordings towards stable, long-term brain-computer interfaces. *Communications Biology*, 7(1):1–14, October 2024. Publisher: Nature Publishing Group.

[59] Samuel R. Nason, Matthew J. Mender, Alex K. Vaskov, Matthew S. Willsey, Nishant Ganesh Kumar, Theodore A. Kung, Parag G. Patil, and Cynthia A. Chestek. Real-time linear prediction of simultaneous and independent movements of two finger groups using an intracortical brain-machine interface. 109(19):3164–3177.e8.

[60] Vikash Gilja, Paul Nuyujukian, Cindy A. Chestek, John P. Cunningham, Byron M. Yu, Joline M. Fan, Mark M. Churchland, Matthew T. Kaufman, Jonathan C. Kao, Stephen I. Ryu, and Krishna V. Shenoy. A high-performance neural prosthesis enabled by control algorithm design. *Nature Neuroscience*, 15(12):1752–1757, December 2012. Publisher: Nat Neurosci.

[61] Z T Irwin, K E Schroeder, P P Vu, A J Bullard, D M Tat, C S Nu, A Vaskov, S R Nason, D E Thompson, J N Bentley, P G Patil, and C A Chestek. Neural control of finger movement via intracortical brain–machine interface. 14(6):066004. Publisher: IOP Publishing.

[62] Alex K Vaskov, Zachary T Irwin, Samuel R Nason, Philip P Vu, Chrono S Nu, Autumn J Bullard, Mackenna Hill, Naia North, Parag G Patil, and Cynthia A Chestek. Cortical decoding of individual finger group motions using ReFIT kalman filter. 12:751. Publisher: Frontiers Media S.A.

[63] Samuel R. Nason, Alex K. Vaskov, Matthew S. Willsey, Elissa J. Welle, Hyochan An, Philip P. Vu, Autumn J. Bullard, Chrono S. Nu, Jonathan C. Kao, Krishna V. Shenoy, Taekwang Jang, Hun Seok Kim, David Blaauw, Parag G. Patil, and Cynthia A. Chestek. A low-power band of neuronal spiking activity dominated by local single units improves the performance of brain–machine interfaces. pages 1–11. Publisher: Nature Research.

[64] Eran Stark and Moshe Abeles. Predicting Movement from Multiunit Activity. *Journal of Neuroscience*, 27(31):8387–8394, August 2007. Publisher: Society for Neuroscience Section: Articles.

[65] J. Schlag and R. Balvin. Background activity in the cerebral cortex and reticular formation in relation with the electroencephalogram. *Experimental Neurology*, 8(3):203–219, September 1963.

[66] Oliver Rübel, Andrew Tritt, Ryan Ly, Benjamin K Dichter, Satrajit Ghosh, Lawrence Niu, Pamela Baker, Ivan Soltesz, Lydia Ng, Karel Svoboda, Loren Frank, and Kristofer E Bouchard. The Neurodata Without Borders ecosystem for neurophysiological data science. *eLife*, 11:e78362, October 2022. Publisher: eLife Sciences Publications, Ltd.

[67] Krzysztof J. Gorgolewski, Tibor Auer, Vince D. Calhoun, R. Cameron Craddock, Samir Das, Eugene P. Duff, Guillaume Flandin, Satrajit S. Ghosh, Tristan Glatard, Yaroslav O. Halchenko, Daniel A. Handwerker, Michael Hanke, David Keator, Xiangrui Li, Zachary Michael, Camille Maumet, B. Nolan Nichols, Thomas E. Nichols, John Pellman, Jean-Baptiste Poline, Ariel Rokem, Gunnar Schaefer, Vanessa Sochat, William Triplett, Jessica A. Turner, Gaël Varoquaux, and Russell A. Poldrack. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific Data*, 3(1):160044, June 2016. Publisher: Nature Publishing Group.

[68] Peiran Gao, Eric Trautmann, Byron Yu, Gopal Santhanam, Stephen Ryu, Krishna Shenoy, and Surya Ganguli. A theory of multineuronal dimensionality, dynamics and measurement. Pages: 214262 Section: New Results.

[69] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997. Publisher: MIT press.

[70] Eric Y. Wang, Paul G. Fahey, Zhuokun Ding, Stelios Papadopoulos, Kayla Ponder, Marissa A. Weis, Andersen Chang, Taliah Muhammad, Saumil Patel, Zhiwei Ding, Dat Tran, Jiakun Fu, Casey M. Schneider-Mizell, R. Clay Reid, Forrest Collman, Nuno Maçarico da Costa, Katrin Franke, Alexander S. Ecker, Jacob Reimer, Xaq Pitkow, Fabian H. Sinz, and Andreas S. Tolias. Foundation model of neural activity predicts response to new stimulus types and anatomy, August 2024. Pages: 2023.03.21.533548 Section: New Results.

[71] Ivan Ustyuzhaninov, Max F. Burg, Santiago A. Cadena, Jiakun Fu, Taliah Muhammad, Kayla Ponder, Emmanouil Froudarakis, Zhiwei Ding, Matthias Bethge, Andreas S. Tolias, and Alexander S. Ecker. Digital twin reveals combinatorial code of non-linear computations in the mouse primary visual cortex, February 2022. Pages: 2022.02.10.479884 Section: New Results.

[72] Cátia Fortunato, Jorge Bennasar-Vázquez, Junchol Park, Joanna C. Chang, Lee E. Miller, Joshua T. Dudman, Matthew G. Perich, and Juan A. Gallego. Nonlinear manifolds underlie neural population activity during behaviour, April 2024. Pages: 2023.07.18.549575 Section: New Results.

[73] Mukund Balasubramanian and Eric L. Schwartz. The Isomap Algorithm and Topological Stability. *Science*, 295(5552):7–7, January 2002. Publisher: American Association for the Advancement of Science.

[74] Jean-Alban Rathelot and Peter L. Strick. Subdivisions of primary motor cortex based on cortico-motoneuronal cells. *Proceedings of the National Academy of Sciences*, 106(3):918–923, January 2009. Publisher: Proceedings of the National Academy of Sciences.

NeurIPS Paper Checklist

1. **Claims**

   Question: Do the main claims made in the abstract and introduction accurately reflect the paper's contributions and scope?

   Answer: [Yes]

   Justification: We present a large dataset of non-human primate intracortical and behavioral data, and analyses of the relationship between the neural data and kinematics as well as decoding performance, as stated in the abstract and introduction. The dataset is described in Section 3 and the analyses are presented in Section 4.

   Guidelines:

   - The answer NA means that the abstract and introduction do not include the claims made in the paper.
   - The abstract and/or introduction should clearly state the claims made, including the contributions made in the paper and important assumptions and limitations. A No or NA answer to this question will not be perceived well by the reviewers.
   - The claims made should match theoretical and experimental results, and reflect how much the results can be expected to generalize to other settings.
   - It is fine to include aspirational goals as motivation as long as it is clear that these goals are not attained by the paper.

2. **Limitations**

   Question: Does the paper discuss the limitations of the work performed by the authors?

   Answer: [Yes]

   Justification: In Section 5, we include a paragraph discussing the limitations of the dataset.

   Guidelines:

   - The answer NA means that the paper has no limitation while the answer No means that the paper has limitations, but those are not discussed in the paper.
   - The authors are encouraged to create a separate "Limitations" section in their paper.
   - The paper should point out any strong assumptions and how robust the results are to violations of these assumptions (e.g., independence assumptions, noiseless settings, model well-specification, asymptotic approximations only holding locally). The authors should reflect on how these assumptions might be violated in practice and what the implications would be.
   - The authors should reflect on the scope of the claims made, e.g., if the approach was only tested on a few datasets or with a few runs. In general, empirical results often depend on implicit assumptions, which should be articulated.
   - The authors should reflect on the factors that influence the performance of the approach. For example, a facial recognition algorithm may perform poorly when image resolution is low or images are taken in low lighting. Or a speech-to-text system might not be used reliably to provide closed captions for online lectures because it fails to handle technical jargon.
   - The authors should discuss the computational efficiency of the proposed algorithms and how they scale with dataset size.
   - If applicable, the authors should discuss possible limitations of their approach to address problems of privacy and fairness.
   - While the authors might fear that complete honesty about limitations might be used by reviewers as grounds for rejection, a worse outcome might be that reviewers discover limitations that aren't acknowledged in the paper. The authors should use their best judgment and recognize that individual actions in favor of transparency play an important role in developing norms that preserve the integrity of the community. Reviewers will be specifically instructed to not penalize honesty concerning limitations.

3. **Theory assumptions and proofs**

   Question: For each theoretical result, does the paper provide the full set of assumptions and a complete (and correct) proof?

Answer: [NA]

Justification: Our paper does not include theoretical results.

Guidelines:

- The answer NA means that the paper does not include theoretical results.
- All the theorems, formulas, and proofs in the paper should be numbered and cross-referenced.
- All assumptions should be clearly stated or referenced in the statement of any theorems.
- The proofs can either appear in the main paper or the supplemental material, but if they appear in the supplemental material, the authors are encouraged to provide a short proof sketch to provide intuition.
- Inversely, any informal proof provided in the core of the paper should be complemented by formal proofs provided in appendix or supplemental material.
- Theorems and Lemmas that the proof relies upon should be properly referenced.

4. **Experimental result reproducibility**

Question: Does the paper fully disclose all the information needed to reproduce the main experimental results of the paper to the extent that it affects the main claims and/or conclusions of the paper (regardless of whether the code and data are provided or not)?

Answer: [Yes]

Justification: In Section 3, we provide a link to the publicly available github repository containing all the code needed to reproduce the results shown in Section 4.

Guidelines:

- The answer NA means that the paper does not include experiments.
- If the paper includes experiments, a No answer to this question will not be perceived well by the reviewers: Making the paper reproducible is important, regardless of whether the code and data are provided or not.
- If the contribution is a dataset and/or model, the authors should describe the steps taken to make their results reproducible or verifiable.
- Depending on the contribution, reproducibility can be accomplished in various ways. For example, if the contribution is a novel architecture, describing the architecture fully might suffice, or if the contribution is a specific model and empirical evaluation, it may be necessary to either make it possible for others to replicate the model with the same dataset, or provide access to the model. In general. releasing code and data is often one good way to accomplish this, but reproducibility can also be provided via detailed instructions for how to replicate the results, access to a hosted model (e.g., in the case of a large language model), releasing of a model checkpoint, or other means that are appropriate to the research performed.
- While NeurIPS does not require releasing code, the conference does require all submissions to provide some reasonable avenue for reproducibility, which may depend on the nature of the contribution. For example
  (a) If the contribution is primarily a new algorithm, the paper should make it clear how to reproduce that algorithm.
  (b) If the contribution is primarily a new model architecture, the paper should describe the architecture clearly and fully.
  (c) If the contribution is a new model (e.g., a large language model), then there should either be a way to access this model for reproducing the results or a way to reproduce the model (e.g., with an open-source dataset or instructions for how to construct the dataset).
  (d) We recognize that reproducibility may be tricky in some cases, in which case authors are welcome to describe the particular way they provide for reproducibility. In the case of closed-source models, it may be that access to the model is limited in some way (e.g., to registered users), but it should be possible for other researchers to have some path to reproducing or verifying the results.

5. **Open access to data and code**

Question: Does the paper provide open access to the data and code, with sufficient instructions to faithfully reproduce the main experimental results, as described in supplemental material?

Answer: [Yes]

Justification: As this is a dataset release paper, the data is publicly available at the link posted in Section 3. Additionally, the GitHub repository with all the code needed to reproduce all results in Section 4 is also included in Section 3.

Guidelines:

- The answer NA means that paper does not include experiments requiring code.
- Please see the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- While we encourage the release of code and data, we understand that this might not be possible, so "No" is an acceptable answer. Papers cannot be rejected simply for not including code, unless this is central to the contribution (e.g., for a new open-source benchmark).
- The instructions should contain the exact command and environment needed to run to reproduce the results. See the NeurIPS code and data submission guidelines (`https://nips.cc/public/guides/CodeSubmissionPolicy`) for more details.
- The authors should provide instructions on data access and preparation, including how to access the raw data, preprocessed data, intermediate data, and generated data, etc.
- The authors should provide scripts to reproduce all experimental results for the new proposed method and baselines. If only a subset of experiments are reproducible, they should state which ones are omitted from the script and why.
- At submission time, to preserve anonymity, the authors should release anonymized versions (if applicable).
- Providing as much information as possible in supplemental material (appended to the paper) is recommended, but including URLs to data and code is permitted.

6. **Experimental setting/details**

Question: Does the paper specify all the training and test details (e.g., data splits, hyperparameters, how they were chosen, type of optimizer, etc.) necessary to understand the results?

Answer: [Yes]

Justification: Training and test details are mostly specified in the Supplemental materials, with any further details included in the code (which is also publicly available and referenced in the paper).

Guidelines:

- The answer NA means that the paper does not include experiments.
- The experimental setting should be presented in the core of the paper to a level of detail that is necessary to appreciate the results and make sense of them.
- The full details can be provided either with the code, in appendix, or as supplemental material.

7. **Experiment statistical significance**

Question: Does the paper report error bars suitably and correctly defined or other appropriate information about the statistical significance of the experiments?

Answer: [Yes]

Justification: Yes, where appropriate, error bars are reported, such as in Figure 2D and Figure 4D.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The authors should answer "Yes" if the results are accompanied by error bars, confidence intervals, or statistical significance tests, at least for the experiments that support the main claims of the paper.

- The factors of variability that the error bars are capturing should be clearly stated (for example, train/test split, initialization, random drawing of some parameter, or overall run with given experimental conditions).
- The method for calculating the error bars should be explained (closed form formula, call to a library function, bootstrap, etc.)
- The assumptions made should be given (e.g., Normally distributed errors).
- It should be clear whether the error bar is the standard deviation or the standard error of the mean.
- It is OK to report 1-sigma error bars, but one should state it. The authors should preferably report a 2-sigma error bar than state that they have a 96% CI, if the hypothesis of Normality of errors is not verified.
- For asymmetric distributions, the authors should be careful not to show in tables or figures symmetric error bars that would yield results that are out of range (e.g. negative error rates).
- If error bars are reported in tables or plots, The authors should explain in the text how they were calculated and reference the corresponding figures or tables in the text.

8. **Experiments compute resources**

Question: For each experiment, does the paper provide sufficient information on the computer resources (type of compute workers, memory, time of execution) needed to reproduce the experiments?

Answer: [Yes]

Justification: The computer resources needed to reproduce the experiments are described in Appendix C of the Supplemental materials.

Guidelines:

- The answer NA means that the paper does not include experiments.
- The paper should indicate the type of compute workers CPU or GPU, internal cluster, or cloud provider, including relevant memory and storage.
- The paper should provide the amount of compute required for each of the individual experimental runs as well as estimate the total compute.
- The paper should disclose whether the full research project required more compute than the experiments reported in the paper (e.g., preliminary or failed experiments that didn't make it into the paper).

9. **Code of ethics**

Question: Does the research conducted in the paper conform, in every respect, with the NeurIPS Code of Ethics `https://neurips.cc/public/EthicsGuidelines`?

Answer: [Yes]

Justification: With respect to the NeurIPS Code of Ethics, this paper conforms in every respect. With regards to Representative evaluation practice, we make no claims that this dataset represents all intracortical neural datasets, with regard to Data and model documentation and licensing, all data are structured according to the NWB standard, valid Croissant metadata is provided, and both data and code are licensed (licenses are available at their respective URLs).

Guidelines:

- The answer NA means that the authors have not reviewed the NeurIPS Code of Ethics.
- If the authors answer No, they should explain the special circumstances that require a deviation from the Code of Ethics.
- The authors should make sure to preserve anonymity (e.g., if there is a special consideration due to laws or regulations in their jurisdiction).

10. **Broader impacts**

Question: Does the paper discuss both potential positive societal impacts and negative societal impacts of the work performed?

Answer: [Yes]

Justification: As discussed in Sections 1 and 5, the release of this dataset is directly motivated by a desire to develop better brain-machine interfaces and improve the lives of those living with paralysis. We do not see a direct path to a negative societal impact with this dataset release.

Guidelines:

- The answer NA means that there is no societal impact of the work performed.
- If the authors answer NA or No, they should explain why their work has no societal impact or why the paper does not address societal impact.
- Examples of negative societal impacts include potential malicious or unintended uses (e.g., disinformation, generating fake profiles, surveillance), fairness considerations (e.g., deployment of technologies that could make decisions that unfairly impact specific groups), privacy considerations, and security considerations.
- The conference expects that many papers will be foundational research and not tied to particular applications, let alone deployments. However, if there is a direct path to any negative applications, the authors should point it out. For example, it is legitimate to point out that an improvement in the quality of generative models could be used to generate deepfakes for disinformation. On the other hand, it is not needed to point out that a generic algorithm for optimizing neural networks could enable people to train models that generate Deepfakes faster.
- The authors should consider possible harms that could arise when the technology is being used as intended and functioning correctly, harms that could arise when the technology is being used as intended but gives incorrect results, and harms following from (intentional or unintentional) misuse of the technology.
- If there are negative societal impacts, the authors could also discuss possible mitigation strategies (e.g., gated release of models, providing defenses in addition to attacks, mechanisms for monitoring misuse, mechanisms to monitor how a system learns from feedback over time, improving the efficiency and accessibility of ML).

11. **Safeguards**

Question: Does the paper describe safeguards that have been put in place for responsible release of data or models that have a high risk for misuse (e.g., pretrained language models, image generators, or scraped datasets)?

Answer: [NA]

Justification: As this paper contains a non-human primate neural dataset, we do not believe it poses high risk for misuse.

Guidelines:

- The answer NA means that the paper poses no such risks.
- Released models that have a high risk for misuse or dual-use should be released with necessary safeguards to allow for controlled use of the model, for example by requiring that users adhere to usage guidelines or restrictions to access the model or implementing safety filters.
- Datasets that have been scraped from the Internet could pose safety risks. The authors should describe how they avoided releasing unsafe images.
- We recognize that providing effective safeguards is challenging, and many papers do not require this, but we encourage authors to take this into account and make a best faith effort.

12. **Licenses for existing assets**

Question: Are the creators or original owners of assets (e.g., code, data, models), used in the paper, properly credited and are the license and terms of use explicitly mentioned and properly respected?

Answer: [Yes]

Justification: The dataset and code were produced by the authors of the publication, and where relevant, citations to previous publications are included.

Guidelines:

- The answer NA means that the paper does not use existing assets.
- The authors should cite the original paper that produced the code package or dataset.
- The authors should state which version of the asset is used and, if possible, include a URL.
- The name of the license (e.g., CC-BY 4.0) should be included for each asset.
- For scraped data from a particular source (e.g., website), the copyright and terms of service of that source should be provided.
- If assets are released, the license, copyright information, and terms of use in the package should be provided. For popular datasets, `paperswithcode.com/datasets` has curated licenses for some datasets. Their licensing guide can help determine the license of a dataset.
- For existing datasets that are re-packaged, both the original license and the license of the derived asset (if it has changed) should be provided.
- If this information is not available online, the authors are encouraged to reach out to the asset's creators.

13. **New assets**

Question: Are new assets introduced in the paper well documented and is the documentation provided alongside the assets?

Answer: [Yes]

Justification: The core of the paper, the LINK dataset is described in Section 3 as well as in further detail in Appendix A and B. All documentation is provided here or with the dataset and code.

Guidelines:

- The answer NA means that the paper does not release new assets.
- Researchers should communicate the details of the dataset/code/model as part of their submissions via structured templates. This includes details about training, license, limitations, etc.
- The paper should discuss whether and how consent was obtained from people whose asset is used.
- At submission time, remember to anonymize your assets (if applicable). You can either create an anonymized URL or include an anonymized zip file.

14. **Crowdsourcing and research with human subjects**

Question: For crowdsourcing experiments and research with human subjects, does the paper include the full text of instructions given to participants and screenshots, if applicable, as well as details about compensation (if any)?

Answer: [NA]

Justification: This paper does not contain crowdsourcing experiments or research with human subjects

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Including this information in the supplemental material is fine, but if the main contribution of the paper involves human subjects, then as much detail as possible should be included in the main paper.
- According to the NeurIPS Code of Ethics, workers involved in data collection, curation, or other labor should be paid at least the minimum wage in the country of the data collector.

15. **Institutional review board (IRB) approvals or equivalent for research with human subjects**

Question: Does the paper describe potential risks incurred by study participants, whether such risks were disclosed to the subjects, and whether Institutional Review Board (IRB) approvals (or an equivalent approval/review based on the requirements of your country or institution) were obtained?

Answer: [NA]

Justification: This paper does not include any human subjects.

Guidelines:

- The answer NA means that the paper does not involve crowdsourcing nor research with human subjects.
- Depending on the country in which research is conducted, IRB approval (or equivalent) may be required for any human subjects research. If you obtained IRB approval, you should clearly state this in the paper.
- We recognize that the procedures for this may vary significantly between institutions and locations, and we expect authors to adhere to the NeurIPS Code of Ethics and the guidelines for their institution.
- For initial submissions, do not include any information that would break anonymity (if applicable), such as the institution conducting the review.

16. **Declaration of LLM usage**

Question: Does the paper describe the usage of LLMs if it is an important, original, or non-standard component of the core methods in this research? Note that if the LLM is used only for writing, editing, or formatting purposes and does not impact the core methodology, scientific rigorousness, or originality of the research, declaration is not required.

Answer: [NA]

Justification: LLMs were not involved in the core method development of this paper.

Guidelines:

- The answer NA means that the core method development in this research does not involve LLMs as any important, original, or non-standard components.
- Please refer to our LLM policy (`https://neurips.cc/Conferences/2025/LLM`) for what should or should not be described.