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], Aren I. Hite[1], Dylan M. Wallace[5], Samuel R. Nason-Tomaszewski[1], Parag G. Patil[1,6], Matt S. Willsey[1,6], Anne W. Draelos[1,7], Cynthia Chestek[1,5]

[1]Biomedical Engineering, [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

BIOMEDICAL ENGINEERING
UNIVERSITY OF MICHIGAN

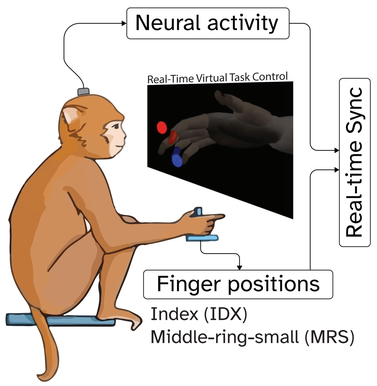
Introduction

- Intracortical brain-machine interfaces limited by signal instability requiring frequent recalibration [1]
- Approaches include recalibrating decoders or aligning neural data [2, 3].
- Few datasets for benchmarking these methods on long time scales [4]

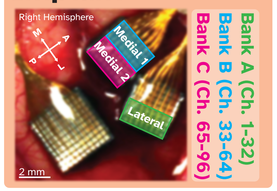
Dataset overview

- Rhesus macaque performed **2-DOF,** trial-based individuated **finger movement task**.
- LINK contains: **312 sessions** on 303 days, spanning 1,242 total days (**~3.5 years**).
- In each session: **375 trials per session, 96 neural channels** preprocessed into **2 neural features**, **position and velocity** for both DOF. 20ms bin size. **117,000 total trials.** Each session is one of two task variations.
- Data hosted on **DANDI**, each session is a separate .nwb (Neurodata without Borders) file.

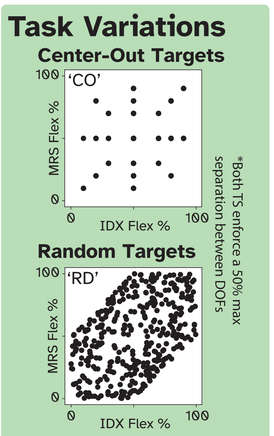
Experimental Setup



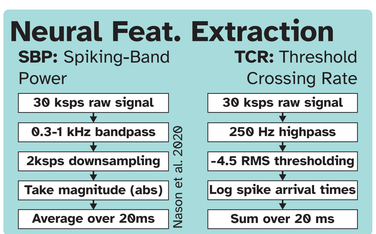
Implant



Task Variations



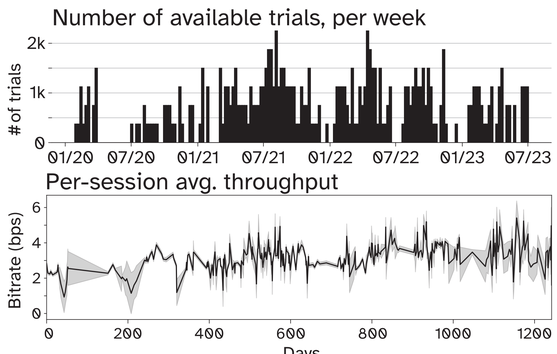
Neural Feat. Extraction



Top Left: Experimental Setup/Behavioral Task. Monkey N performed a 2-degree-of-freedom (DOF) trial-based finger movement task, moving two finger, index and middle-ring-small in a custom manipulandum (with flex sensors for motion tracking) to acquire spherical targets presented on a screen. Monkey N received a juice reward upon successfully acquiring targets. **Top Right:** Monkey N was implanted with 3 Utah arrays. Neural data in the LINK dataset comes from 96 channels highlighted in the figure, from the hand area of the precentral gyrus/motor cortex. **Bottom Left:** Neural data in the LINK dataset was preprocessed into 2 neural features, spiking-band power (SBP) and threshold crossing rate (TCR), and pre-binned into 20ms timepoints. **Bottom Right:** Targets for contiguous sets of trials (dubbed sessions) were presented in one of two task variations, 'center out' (CO), where every other trial is a return to a resting position, and 'random targets' (RD), where targets are selected pseudo-randomly. In all cases, split targets were limited to 50% of the total movement range for comfort.

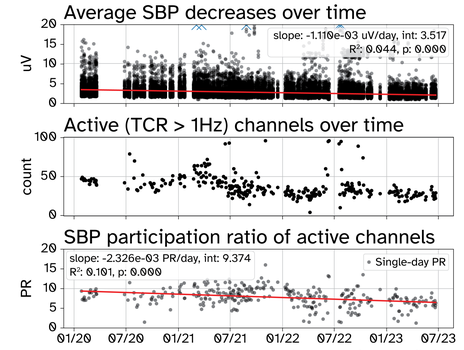
Data characterization

Trials available and behavior



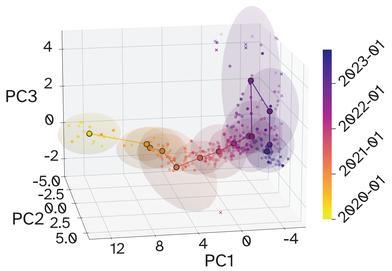
Top: Histogram of trial availability across the dataset. Bins are a week long. **Bottom:** Average performance on behavioral trial. Each point on the line is the average throughput (Fitt's Law) in bits/s across all trials from one day.

Shifts in neural data over time

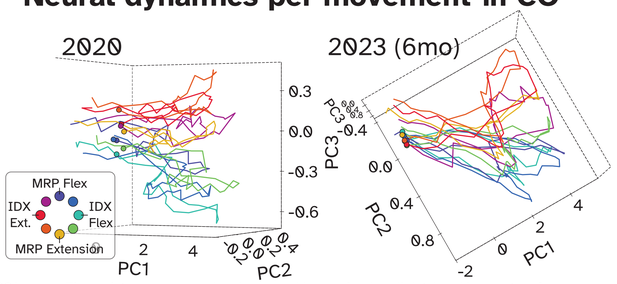


Top: Each dot is the average SBP of one channel in one session. Best-fit line shown in red. **Middle:** Number of active channels (defined as channels with avg. TCR > 1 Hz across a session). **Top:** Participation ratio [5] (analog for dimensionality) of the SBP of active channels over the dataset. Best-fit line shown in red.

SBP drifts in full-data PCA space



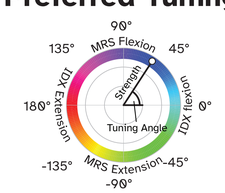
Neural dynamics per movement in CO



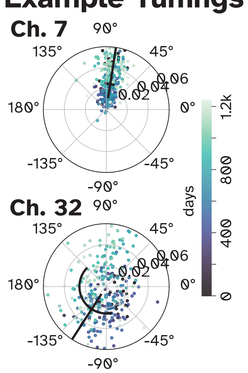
Left: SBP data from all sessions was concatenated and projected into a full-data PCA space. Timepoints were grouped and averaged by session (small dots) and by quarter of a year (large dots) and plotted in top-3 PC space, shown above. SD of per-quarter groups are shown using the shaded ellipsoids. **Right:** SBP data from center-out sessions were projected into within session PC spaces. Trials starting from 'rest' were grouped by movement direction and by year, aligned by movement onset, and averaged in PC space. The resulting average neural dynamics are plotted above, with color indicating movement direction.

Single channel tuning to behavior over time

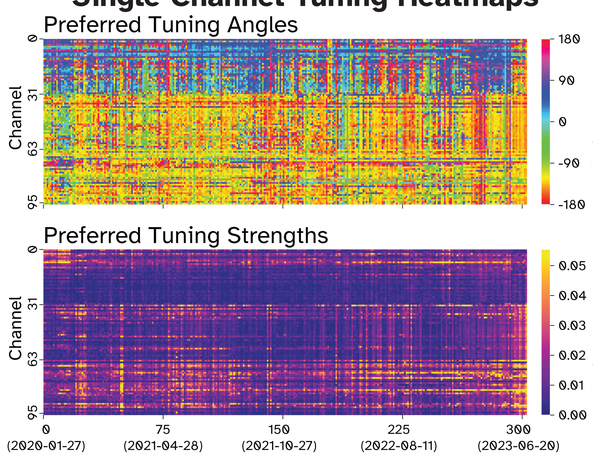
Preferred Tuning



Example Tunings

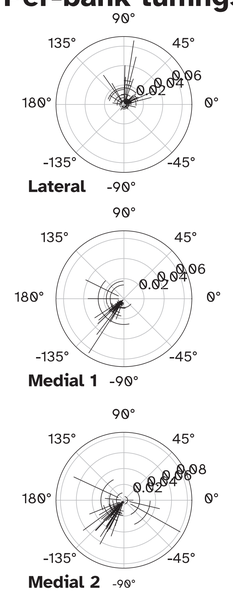


Single Channel Tuning Heatmaps



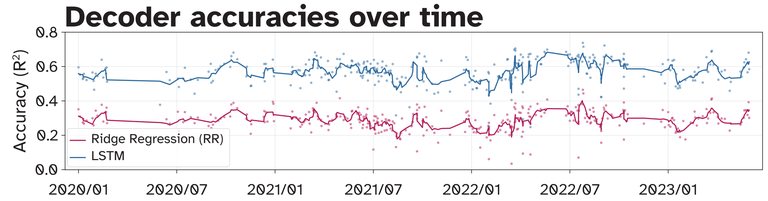
Top Left: Toy example of preferred tuning for a single neural channel. Preferred tuning was measured by stacking the coefficients of linear regressions to each degree of freedom, for a single neural channel (in a single single) **Bottom Left:** Example tunings for two channels across the whole dataset. Each dot is a tuning measurement for one session, with color indicating the day. Black bars show the median and IQR of tuning across the dataset (circular for tuning angle) **Middle:** Tuning angles and strengths for each channel on each session, presented as two heatmaps. **Right:** Median and IQR of tunings a cross dataset for each channel, split into 32 channel groups. Circular median and IQR used for angles.

Per-bank tunings

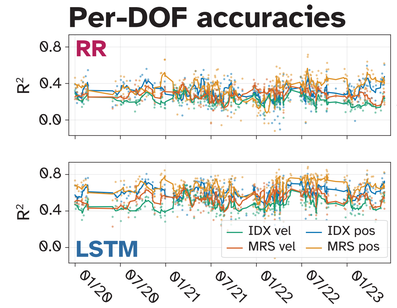


Decoding accuracies over time

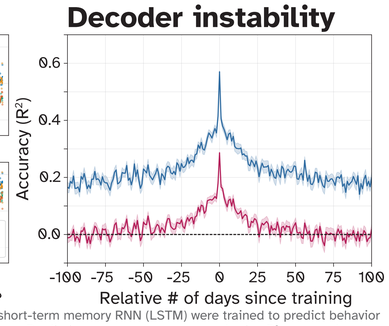
Decoder accuracies over time



Per-DOF accuracies



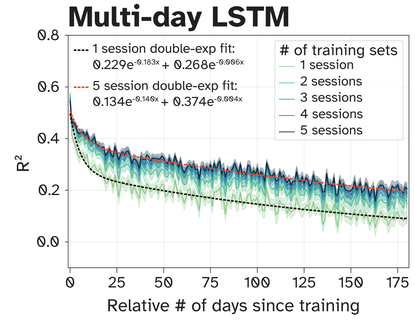
Decoder instability



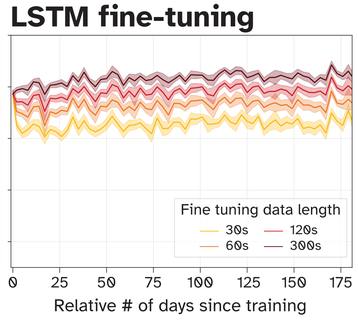
Top: Within each session, a Ridge Regression (RR) and a long short-term memory RNN (LSTM) were trained to predict behavior from neural data using 300 trials and evaluated on 75 trial holdouts. Prediction accuracy was measured using R^2, averaged across DOFS (position and velocity). The two lines smooth the performances for visual clarity. **Bottom Left:** R^2 for each DOF across the dataset, for each decoder. Dots show real performance, lines are smoothed for clarity. **Bottom Right:** All decoders were then tested on the holdout sets for all sessions, and average performance was labelled by the relative amount of time since the training session. Performances were averaged by relative day and plotted here. Shaded areas are +/- SE.

Preliminary recalibration

Multi-day LSTM



LSTM fine-tuning



Left: LSTMs were trained on 300 trials from between 1-5 successive datasets, and tested on the holdouts of subseuent datasets. We once again averaged by relative day from the latest dataset included in the training data. Line color indicates the number of sessionused for training. We observed that decoding per appeared to folow an exponential decay, so we fit double-exponentials to the 1- and 5-session performances, shown in the black and red lines, respectively. Shaded areas are +/- SE. **Right:** Here, LSTMs pre-trained on single sessions, and fine-tuned on limited amounts of data on all datasets up to the relative day on the plot. Here, 20 seed days were used as the initial single sessions. LSTMs were fine tuned on varying amounts of data, as indicated by the line colors. PErformance seems to increase over the dataset (after an initial drop, in some cases), likely because LSTMs were fine-tuned on all subsequent datasets leading to the relative day on the plot.

Acknowledgements

A massive thank you to Monkey N, who's contributions to the field have been instrumental in advancing brain machine interfaces, to Eric Kennedy, to the Draelos Lab, and to the Chestek Lab.

Funding

- National Science Foundation Grant 1926576 (NCS)
- National Science Foundation Grant 2223822 (EFRI/BRAID)
- National Institutes of Health Grant T32NS007222
- National Institutes of Health Grant AWD027941
- National Science Foundation GRFP Grant 1841052
- Agencia Nacional de Investigación y Desarrollo (ANID) of Chile
- Dan and Betty Kahn Foundation Grant 2029755
- Michigan Robotics Institute
- CASI Award #1021865.01 from the Burroughs Wellcome Fund

Citations:

1. Gallego, JA., et al. 2020. Nat Neuro
2. Degenhart AD., et al. 2020. Nat Biom Eng
3. Fan, C., et al. 2023. NeurIPS
4. Karpowicz et al. 2024 NeurIPS
5. Gao et al. 2017, biorxiv

NSF
NIH
National Institutes of Health