# Local network coordination supports neuroprosthetic control

William Liberti III[1], Xue Lily Gong[2], Thomas Roseberry[3], and Jose Carmena[4], *Senior Member, IEEE*

*Abstract*— **Learning often involves adapting behavior in response to the inferred causes of success and failure. At the neural level, this can be the result of repeating activity patterns of neurons that lead to favorable outcomes. However, it is not clear how networks assess the contributions of individual cells to an ongoing behavior. Using a calcium imaging based closed loop Brain Machine Interface (BMI), we trained mice to perform a neuroprosthetic task using the coordinated activity of a small ensemble of neurons in layer 2/3 of sensorimotor cortex. Throughout a single 30 minute session, network conformations that result in reward dramatically increase in frequency relative to similar conformations in other cells. We find that that after an initial period of exploration, neurons unrelated to the task demonstrate a decrease in variance and event frequency over the course of learning. However, many of these 'indirect' cells demonstrate robust spatiotemporal dynamics both before and after an animal achieves reward. We estimate the relative contributions of the local network by using nearby, unrelated cells to decode the activity of the BMI neurons compared to arbitrary neurons uninvolved with this task. Taken together, our findings suggest that neuroprosthetic control of output cells is the result of specific spatiotemporal ensemble patterns that are quickly identified and reinforced.**

## I. INTRODUCTION

Several studies of learning have shown that spatiotemporal activity patterns of neurons which lead to desired behavioral outcomes are selected and consolidated [1]–[3]. The theoretical framework that describes this underlying computational problem is called 'credit assignment', and it is an important ingredient of reinforcement learning [4]. However, the single neuron rules that result in behavior modification or maintenance have been difficult to address, due to the immense fine-scale complexity of biological neural networks and the heterogeneity of natural behavior. Traditional behavioral assays are agnostic to the underlying neural dynamics used by animals to accomplish a given task, making it nearly impossible to assert the causal role of individual neurons to a behavioral action. Moreover, natural tasks are thought to engage hundreds of thousands of neurons; most of these elements are unobservable to the experimenter, making it impossible to estimate the redundancy and degeneracy of neural codes that coordinate learned skills. With current technology, a holistic understanding of the single unit code

that underlies natural skill learning is an under-constrained problem.

To overcome these obstacles, we employ a reductionist approach where the behavioral output is simply the optically recorded activity of a small ensemble of experimenter defined cortical neurons expressing genetically encoded fluorescent calcium indicators. This neuroprosthetic strategy uses closed-loop auditory feedback to guide a 'fictive behavior' where animals are rewarded when a specific network conformation is observed in a chosen neural population [5], [6]. Similar to natural motor tasks, previous work has shown that neuroprosthetic task performance tends to increases over days [5]. Surprisingly, BMI performance can increase even when the output cell identity is changed from day to day- suggesting that mice are learning to better assign reward credit to output neurons, rather than consolidating specific activity control patterns [7]–[9]. It is unclear how the local population of cells resolves the identity of the output cells on these short timescales. The remaining units that are not assigned a direct relationship with effector, called *'indirect neurons'*, do demonstrate some tuning relative to BMI tasks, suggesting that some these neurons also become incorporated into functional neuronal assemblies that help coordinate neurons that directly drive output [10], [11]. However, limited cell yields from chronic electrophysiology and an experimenter bias to assign the highest signal-to-noise units as output cells has provided limited opportunities to ask questions about the relative contributions of local networks [11].

After an initial period of exploration, we find that the local cortical neural population converges to form reproducible, spatiotemporally organized activity patterns that increases in frequency and robustness through a 30 minute session [5], [10]. Our findings suggest that neuroprosthetic control can be the result of an emergent, spatially organized network level solution.

## II. METHODS

### A. Animals and Surgical Procedure.

All animal procedures were performed in accordance with U.C. Berkeley IACUC regulations. Stereotactic surgery was performed in (n=5) $Emx1^{IREScre}$ or wild type mice as previously described [5]. 200-300nl of AAV2.9 CamKII.GCamp6f.WPRE.SV40 or AAV5-SYN-FLEX-GCaMP6f-WPRE-SV40 was used to preferentially label excitatory cells in cortex. Mice were given 4 weeks post-surgical recovery time to allow for sufficient protein expression.

[1]William Liberti is a Fellow in the Department of Electrical Engineering and Computer Science, U.C. Berkeley, Berkeley, CA 94708, USA bliberti@bu.edu

[2]Lily Xue Gong is a Graduate Student in the Helen Wills Neuroscience Institute, U.C. Berkeley, Berkeley, CA 94708, USA

[3]Thomas Roseberry is an engineer at Neuralink, San Francisco, CA 94708, USA

[4]Jose Carmena is a full professor in the Department of Electrical Engineering and Computer Science, U.C. Berkeley, Berkeley, CA 94708, USA jcarmena@berkeley.edu
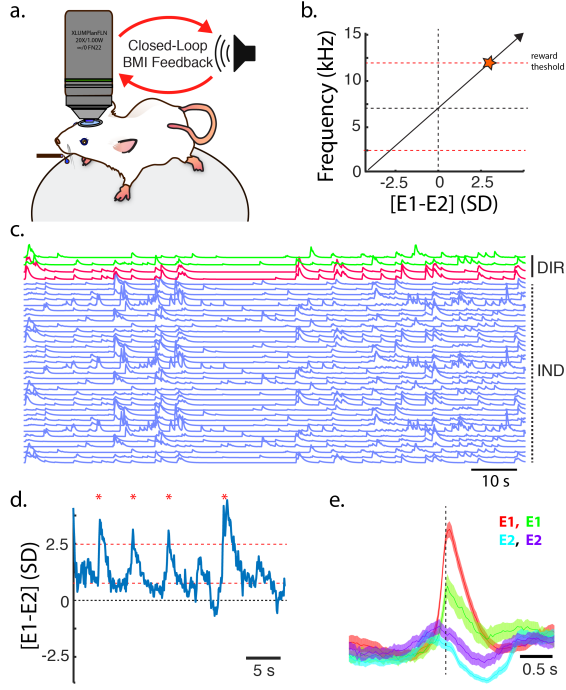
Fig. 1. *Experimental paradigm for a closed-loop calcium imaging based auditory feedback.* ($A$) Head-fixed mice run freely on a spherical treadmill, and perform BMI experiments for a water reward. ($B$) Mice must modulate calcium dynamics in neuronal ensembles to move a cursor to a high-pitched target tone that was associated with reward. ($C$) Two ensembles ( eg, $E1$ in green; $E2$ in red) of 2 single cells each were chosen for inclusion in the 'direct' output population. We were able to observe hundreds of additional 'indirect' neurons in the same field of view, their $\Delta F/F_0$ over time is schematized in blue. ($D$) Mice were able to modulate the calcium dynamics in these neuronal ensembles to move a cursor $2.5 SD$ above baseline levels. Additional reward was prevented until the cursor returned to within baseline levels ( $1SD$ of the mean). ($E$) Output neuron activity converges to form reproducible activity patterns to achieve a reward. Shown here is the average $E1$ and $E2$ population centered at the hit. Shading represents $2\mathrm{x}sem$.

### B. Experimental Preparation

In vivo imaging was performed at 30Hz using a commercial multi-photon microscope (Bruker, Ultima Investigator) driven by a mode-locked, tunable Ti:Sapphire laser (CHAMELEON ULTRA II) and set to 900-980 nm, steered via 8kHz resonant galvos through a 20x water immersion objective (XLUMPLFLN 20XW). Photons were collected with a GaAsP photomultiplier tube (Hamamatsu Model H10770). Mice were put on a restricted water regiment for the duration of the experiment [12]. During the task, mice were head-fixed but allowed to run freely on a circular treadmill, consisting of styrofoam ball suspended by air (PhenoSys, JetBall Virtual Reality System). Non-rigid image registration, ROI extraction, baseline de-trending, and spike deconvolution was performed offline. [13], [14]

### C. Behavioral Task

The boundaries of two ensembles of two single cells were manually defined after a 3-15 minute baseline period of imaging. Care was taken to use well isolated regions of interest (ROIs). Depending on the density of viral labeling,

we were able to observe 300-800 additional neurons in a $500um^2$ field of view. Ensemble activity was measured as the summed, running z-score normalized $\Delta F/F_0$ for each component neuron. This normalization is performed to scale each neuron by its dynamic range, in order to overcome potential differences in viral expression across cells, such that each direct unit would contributed equally to the output. The cursor was smoothed with a running average of 2 - 3 frames. Closed-loop latency was 38ms, with a jitter of 17ms ( 95% confidence). Frequencies used for auditory feedback ranged from 1-20 kHz in quarter-octave increments to match rodent psychophysical discrimination thresholds [5]. Mice must modulate calcium dynamics in these neuronal ensembles to move the cursor to a high-pitched target tone that was associated with water reward, set at 2.5 standard deviations ($SD$) above the baseline (Fig 1b). Additional reward was prevented until the cursor returned to within baseline levels ( $1SD$ of the cursor mean).

### D. L2-regularized Linear Regression Model

Indirect neuron features were derived from the spike deconvolved calcium activity within a 2 second sliding window [timepoints * spike and height features]. The response matrix consisted of the fluorescent change ($\Delta F/F_0$ ) of each direct neuron ($E1_a$, $E1_b$, $E2_a$, $E2_b$), the cursor ($E1 - E2$), and the 2 ensembles $E1$ and $E2$. We separated the indirect neuron feature matrix ($IND$) and direct neuron response matrix ($DN$) into a training dataset ($IDN : 54378 * 1580, DN : 54378 * 7$) and completely withheld validation dataset based at the moment of the 'hit': ($IDN : 112 * 1580, DN : 112*7$). To avoid bias, we fixed the regularization coefficient-determined by cross-validating the regression procedure $5\times$. In each cross-validation iteration, 80% of the training dataset was used to estimate the model weights for each of 10 possible regularization coefficients of the remaining 20%. After the completion of cross-validation, a regularization-performance curve was obtained by averaging the cross-validation sample $R^2$ first across 10 regularization coefficients samples and then across all 7 direct neuron responses, then the best regularization parameter (21.544) was selected. Model estimated weights were used to predict responses of the withheld validation dataset. We repeated the modeling procedure using the features of indirect neurons to predict shuffled direct neurons' response matrix in time (frames) axis. We shuffled the response matrix $1000\times$ to build a null distribution of $R^2$ to obtain $P$ values [15].

### III. RESULTS

### A. Evidence Against Solution Degeneracy

Mice were able to modulate the calcium dynamics in these neuronal ensembles to move the cursor to a high-pitched target tone that was associated with water reward (Fig 1c) [5]. An increase in reward frequency could result from a non-specific, or a multitude of degenerate strategies, instead of specifically controlling individual cells. For example, upon hearing a tone, increased attention or arousal may increase unit excitability, increasing cortical network entropy

compared to a baseline period. This feedback loop would increase the chance level of reward compared to baseline- and appear to be volitional control of individual cells [10].

If this was true, we would expect that both the variance across the network, and the average number of events across individual cells in the interval immediately before the hit criteria is met would remain constant or increase over the course of the session, reflecting a strategy of using numerous state permutations to gain reward. However, we observe the opposite- total variance across the network tends to decrease over time, as does the mean population activity of indirect neurons in the moments leading up to the hit (Fig 2c-d). In contrast, cells that are involved with the cursor increase their activity over the session at the moment of the hit (Fig 2a). Given these results, we reasoned that it would be appropriate to normalize the number of rewards over the session as a function of the number of rewards that would have been achieved if any other pairs of neurons had been driving the cursor. This would provide a control to see if the neural population consolidates a limited number of network confor- mations that earned a reward. Compared to the population, we find there is a strong trend to increase reward related conformations over time, supporting the notion that specific activity patterns are being consolidated on this timescale (Fig 2b).
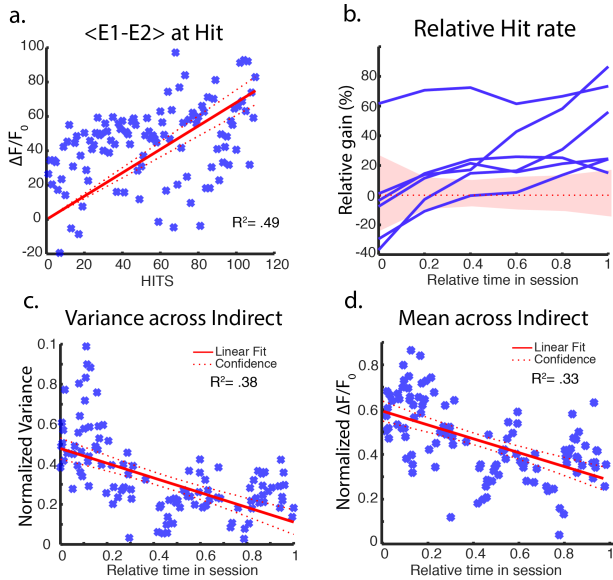


Fig. 2. *Unrewarded neural population activity patterns sparsen over time.* (A) Relative $\Delta F/F_0$ of the output neurons, significantly increases throughout the course of the session ( p<0.01). (B) Compared to the simulated rate of reward using random pairs of indirect neurons, we find there is a strong trend to increase reward related conformations over time. Each blue line represents one animals across a single day of BMI training. The red line indicates the distribution of reward rates for all animals that would have been achieved if any other pairs of neurons had been driving the cursor ( shading is $2 \times sem$). (D) For this same animal in $A$, variance across the population of indirect neurons within one second preceding the hit drops significantly across the session ( p<0.01), (D) as does the mean $\Delta F/F_0$ across the this same interval ( p<0.01). Each blue marker is a sample taken at the 'hit'

## B. Prediction of Output Cells Using Indirect Neuron Activity

Increase of the relative hit rate over time suggests that the network is learning to specifically modulate a subset of neurons to increase the reward rate (Fig 2b). We reasoned that it may be possible to observe network credit assignment by modeling how well the population is able to decode output neuron activity, relative to non-output related cells. Using L2-regularized linear regression (also known as ridge regression), we find that the direct neuron $\Delta F/F_0$ can accurately decoded from the indirect population activity with high accuracy ($R^2$ values for $E1_a$ = 0.77, $E1_b$ = 0.78, $E2_a$ = 0.71, $E2_b$ = 0.53) in addition to the the ensemble activity ( $R^2$ E1 = 0.65, E2 = 0.69). The cursor value was more difficult to decode ($R^2$ = 0.36), suggesting the indirect population encodes the motor output, rather than integrating incoming sensory information. We find that a subset of neurons (10-20%) can decode the majority of the variance in the direct cells. This finding suggests that the network is able to learn to quickly identify and assign credit to specific output neurons to provide neuroprosthetic control. Moreover, we use the weights of the regression model to see which indirect cells care the most about the cursor/direct neurons. Interestingly, we find that these cells are sparsely distributed and not necessarily local to the direct neurons- even though previous reports have shown that indirect cells tend to weakly co-vary with nearby direct neurons, [16], [17].

## C. Spatiotemporal Distribution of Indirect Cells

The fact that a number of indirect neurons can be predicted by the population with similar accuracy as the direct neu- rons, but the relative diversity of network conformations is decreasing throughout the task, suggests that this neuropros- thetic skill may be driven by spatiotemporal ensembles of many cells. Motor and premotor structures have been shown to be locally clustered when driving natural movement, which may result from intrinsic connectivity patterns [18], [19]. To test if this is true in our neuroprosthetic task, for every successful ensemble configuration that satisfied the success criteria of a 'hit', (ie. $E1 - E2 > 2.5SD$), we take the activity 1.5s before and 1.5s after this moment, and observe how many cells have consistent changes in fluorescence across 'hits'. We can compare the mean $\Delta F/F_0$ activity across a subset of hits for each indirect cell, and compared it to the mean ROI activity of an independent subset. ( eg, even vs odd hits). We find that roughly 10-20% this population form a temporal relationship with the direct neurons, and consistently either lead or follow the activity of output neurons(Fig 4a,c) [10]. This result suggests that indirect neurons coordinate the timing of cells that directly drive the effector. Plotting the spatial distribution of these cells reveals striking spatial structure. (Fig 4b,d).

## IV. DISCUSSION

The findings described here largely recapitulate observa- tions seen in across-day learning in both neuroprosthetic, and natural motor tasks: behavioral performance and the consistency neural responses tend to stabilize over time
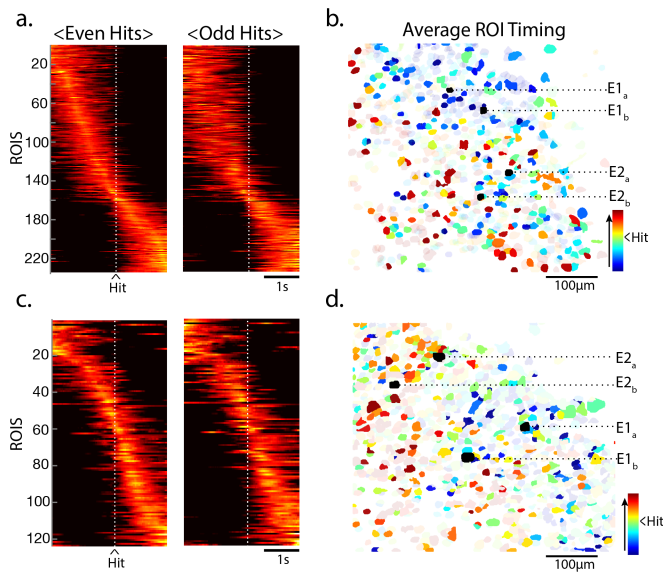
**Fig. 3.** *Indirect neurons form robust spatiotemporal patterns leading up to, and following the rewarded output neuron configuration.* (*A*) Three seconds of activity, centered at the reward criteria ('trials'), were split into two interleaved groups ( 55 each, 110 total for this animal) and averaged across trials. The mean ROI activity was sorted on the even trials, and this sorting is applied to the odd trials . A subset of ROIs demonstrated high consistency ( right panel displays 220/790 identified ROIs that all have >0.95 Pearson correlation). and this correlated with peak amplitude ( p<.01) (*B*) Identified regions of interest colored by time according to the sorting on the even trials in *A*. ROIs with <0.95 correlations are displayed but masked with opacity. (*C* − *D*). Same plot as *A* − *B*, but using data from a previous day of imaging in a nearby cell population, where different output neurons had been selected (displaying 120/500 ROIs, sorted on the even trials from this day).

[8]. In this report, we demonstrate explicit evidence that neuroprosthetic learning within a day can be the result of an initial increase, then rapid decrease of unrewarded neural activity patterns in neural populations that do not control an external effector. This result highlights a classic prediction from reinforcement learning theory: early in a session, there is high variance while the network explores various network conformations- those that lead to reward are consolidated and then exploited, and unrelated activity becomes sparse. It is important to note that neither the variance, or the individual performance seems to continue decreasing by the end of the trial duration ($30min$) suggesting that this may be too short a window to master this particular task. In addition, we observe that the surrounding network can converge in space and time to produce repeated activity that is consistent on average. This suggests that ensembles of neurons, rather than individual cells, can act as the emergent functional unit of neuroprosthetic learning in mice [20]. Possibly, this is the result of re-assigning endogenous existing sequences to achieve reward [21].

The network spatiotemporal organization of motor cortex is unlikely to be predominantly sensory driven, given that neuroprosthetic skills show minimal modulation with passive tone playback [22] In conjunction, we observe that the indirect population cannot decode the tone ($R^2 = 0.36$) as well as the ensemble activity ( mean $R^2 = 0.70$), suggesting that the network can disambiguate reward expectation from reward guided behavior. Important caveats remain. Wile previous work has shown that neuroprosthetic tasks operate without detectable behavior, it possible that motor consequences are sufficiently subtle and go undetected.

## V. CONCLUSION

Neuroprosthetic learning can be particularly insightful in studying the neural basis of sensorimotor integration. By explicitly constraining both the intrinsic jitter found in natural motor production, and the variability of performance evaluation, these experiments allow explicit measurements of how network patterns emerge to form behavior.

### REFERENCES

[1] E. E. Fetz, "Operant conditioning of cortical unit activity," *Science*, vol. 163, no. 3870, pp. 955–958, 1969.

[2] K. Ganguly *et al.*, "Emergence of a stable cortical map for neuroprosthetic control," *PLoS Biol.*, vol. 7, no. 7, p. e1000153, 2009.

[3] V. Athalye *et al.*, "Evidence for a neural law of effect," *Science*, vol. 359, no. 6379, pp. 1024–1029, Mar. 2018.

[4] R. S. Sutton and A. G. Barto, "Reinforcement learning: An introduction," *IEEE Trans. Neural Netw.*, vol. 9, no. 5, pp. 1054–1054, 1998.

[5] K. B. Clancy *et al.*, "Volitional modulation of optically recorded calcium signals during neuroprosthetic learning," *Nat. Neurosci.*, vol. 17, no. 6, pp. 807–809, Jun. 2014.

[6] L. Grosenick *et al.*, "Closed-loop and activity-guided optogenetic control," *Neuron*, vol. 86, no. 1, pp. 106–139, Apr. 2015.

[7] A. Peters *et al.*, "Emergence of reproducible spatiotemporal activity during motor learning," *Nature*, vol. 510, no. 7504, pp. 263–267, Jun. 2014.

[8] T. Gulati *et al.*, "Neural reactivations during sleep determine network credit assignment," *Nat. Neurosci.*, vol. 20, no. 9, pp. 1277–1284, Sep. 2017.

[9] A. Mitani *et al.*, "Brain-Computer interface with inhibitory neurons reveals Subtype-Specific strategies," *Curr. Biol.*, vol. 28, no. 1, pp. 77–83.e4, Jan. 2018.

[10] M. Prsa *et al.*, "Rapid integration of artificial sensory feedback during operant conditioning of motor cortex neurons," *Neuron*, vol. 93, no. 4, pp. 929–939.e6, Feb. 2017.

[11] K. Ganguly *et al.*, "Reversible large-scale modification of cortical networks during neuroprosthetic control," *Nat. Neurosci.*, vol. 14, no. 5, pp. 662–667, May 2011.

[12] Z. V. Guo *et al.*, "Procedures for behavioral experiments in head-fixed mice," *PLoS One*, vol. 9, no. 2, p. e88678, Feb. 2014.

[13] E. A. Pnevmatikakis *et al.*, "NoRMCorre: An online algorithm for piecewise rigid motion correction of calcium imaging data," *J. Neurosci. Methods*, vol. 291, pp. 83–94, Nov. 2017.

[14] ——, "Simultaneous denoising, deconvolution, and demixing of calcium imaging data," *Neuron*, vol. 89, no. 2, pp. 285–299, Jan. 2016.

[15] A. Nunez-Elizalde *et al.*, "Voxelwise encoding models with non-spherical multivariate normal priors," 2018.

[16] B. Sriram *et al.*, "A sparse unreliable distributed code underlies the limits of behavioral discrimination," 2018.

[17] R. Hahnloser *et al.*, "An ultra-sparse code underlies the generation of neural sequences in a songbird," *Nature*, vol. 419, no. 6902, pp. 65–70, Sep. 2002.

[18] J. E. Markowitz and W. A. o. Liberti, "Mesoscopic patterns of neural activity support songbird cortical sequences," *PLoS Biol.*, vol. 13, no. 6, p. e1002158, Jun. 2015.

[19] A. P. Georgopoulos *et al.*, "Mapping of the preferred direction in the motor cortex," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 104, no. 26, pp. 11 068–11 072, Jun. 2007.

[20] R. Yuste, "From the neuron doctrine to neural networks," *Nat. Rev. Neurosci.*, vol. 16, no. 8, pp. 487–497, Aug. 2015.

[21] M. D. Golub *et al.*, "Learning by neural reassociation," *Nat. Neurosci.*, vol. 21, no. 4, pp. 607–616, Apr. 2018.

[22] R. Neely *et al.*, "Volitional modulation of primary visual cortex activity requires the basal ganglia," *Neuron*, vol. 97, no. 6, pp. 1356–1368.e4, Mar. 2018.