Neural Manifold Constraint for Spike Prediction Models Under Behavioral Reinforcement

Shenghui Wu, *Student Member, IEEE*, Xiang Zhang, *Member, IEEE*, and Yiwen Wang, *Senior Member, IEEE*

***Abstract*— Spike prediction models effectively predict downstream spike trains from upstream neural activity for neural prostheses. Such prostheses could potentially restore damaged neural communication pathways using predicted patterns to guide electrical stimulations on downstream. Since the ground truth of downstream neural activity is unavailable for subjects with the damage, reinforcement learning (RL) with behavior-level rewards becomes necessary for model training. However, existing models do not involve any constraint on the generated firing patterns and neglect the correlations among neural activities. Thus, the model outputs can greatly deviate from the natural range of neural activities, causing concerns for clinical usage. This study proposes the neural manifold constraint to solve this problem, shaping RL-generated spike trains in the feature space. The constraint terms describe the first and second order statistics of the neural manifold estimated from neural recordings during subjects' freely moving period. Then, the models can be optimized within the neural manifold by behavioral reinforcement. We test the method to predict primary motor cortex (M1) spikes from medial prefrontal (mPFC) spikes when rats perform the two-lever discrimination task. Results show that the neural activity generated by constrained models resembles the real M1 recordings. Compared with models without constraints, our approach achieves similar behavioral success rates, but reduces the mean squared error of neural firing by 61%. The constraints also increase the model's robustness across data segments and induce realistic neural correlations. Our method provides a promising tool to restore transregional communication with high behavioral performance and more realistic microscopic patterns.**

***Index Terms*— Neural prostheses, neural manifold constraint, neural spike prediction, reinforcement learning.**

Manuscript received 13 March 2024; revised 10 June 2024; accepted 24 July 2024. Date of publication 29 July 2024; date of current version 5 August 2024. This work was supported in part by the STI 2030-Major Projects under Grant 2021ZD0200403, in part by the National Natural Science Foundation of China under Grant 61836003, in part by the Seed Fund of the Big Data for Bio Intelligence Laboratory from The Hong Kong University of Science and Technology (HKUST) under Grant Z0428, in part by the Innovation and Technology Commission under Grant ITCPD/17-9, and in part by the Special Research Support from Chau Hoi Shuen Foundation under Grant R9051. *(Corresponding author: Yiwen Wang.)*

This work involved animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Animal Ethics Committee of the Hong Kong University of Science and Technology (HKUST) IACUC under Application No. AEP-2021-0098.

Shenghui Wu and Xiang Zhang are with the Department of Electronic and Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong (e-mail: shenghui.wu@connect.ust.hk; xzhangaz@connect.ust.hk).

Yiwen Wang is with the Department of Electronic and Computer Engineering and the Department of Chemical and Biological Engineering, The Hong Kong University of Science and Technology, Hong Kong (e-mail: eewangyw@ust.hk).

Digital Object Identifier 10.1109/TNSRE.2024.3435568

I. Introduction

MALFUNCTION of neural pathways among brain regions often results in disorders, such as Alzheimer's disease or spinal cord injuries. Bi-directional neural prostheses offer a potential solution to restore functional loss. These prostheses record upstream neural activity and deliver electrical stimulations to downstream cortical areas [1], [2], [3], [4]. Thus, finding proper stimulation patterns is crucial for modulating the downstream neural activity into desired states and improving behavioral functions. A previous study proposed a predictive model that predicted CA1 neural spike trains as stimulation patterns from CA3 neural spike trains, thereby facilitating human memory recall in behavioral tasks [4]. However, prevailing spike prediction models [4], [5], [6] typically assume that the firing of downstream neurons is independently related to the upstream neural activity. The correlations in the downstream neural population have rarely been considered [7].

The correlation among neurons reflects the intrinsic connectivity of the cortical networks and constrains the distribution of neural firing. Besides, the limitation of single-neuron firing rates also restricts the range of neural activities. As a result, the neural population activity tends to reside within a stable subspace, called the neural manifold [8], [9], [10], [11], [12]. Such a manifold can constrain the adaptation of neural activities during the subject's learning process. Sadtler et al. [9] designed a brain-machine interface (BMI) that mapped the neural population activity of animals into cursor velocities. They found that the subjects quickly learned to control the cursor using neural patterns within an intrinsic low-dimensional subspace, and were less able to learn when required to use neural patterns outside the subspace. These findings suggest that the neural manifold is an essential feature of the neural population activity that must be considered in spike prediction models for cortical stimulation.

The demand for introducing the neural manifold into spike prediction becomes more urgent when the neural pathway is abnormal. Given the disorders, downstream recordings for specific behavioral outputs would be limited or unavailable for supervised learning (SL)-based model training. Recently,

a reinforcement learning (RL) method [13] trained artificial neural networks (ANNs) for spike generation by behavior-level feedback, getting rid of the supervision from target neuron recordings. These RL-trained models outperformed SL-trained models in movement tasks. However, due to the lack of constraints on single-neuron or neural population activities, the model predictions may exhibit very high instantaneous firing rates and quite differ from the natural recordings. These patterns raise concerns for clinical applications, as they may violate the intrinsic neural manifold and hinder the learning process of the subjects. Besides, some inappropriate stimulations may even be unsafe.

Previous research has utilized some constraints on model-generated neural activities. For instance, in clinical trials, researchers restricted the maximum number of stimulations on each electrode channel [4]. In neuroscience, firing rates were regularized to control the power of individual neurons when studying neural dynamics [14], [15]. However, these constraints are too coarse for neural population activity. In fact, neurons can still have different correlations within the same range of firing rates.

In this paper, we propose the neural manifold constraint on RL-based models to generate neural population spike trains from upstream to downstream regions. The generated spiking activity demonstrates realistic neural correlations and leads to behavioral success. Specifically, the constraint terms describe the first and second order statistics of the neural manifold, which is estimated from neural recordings during the freely moving period with principal component analysis (PCA). We derive the constrained policy gradient method for reward-guided model training, which simultaneously restricts the predicted firing probabilities within the neural manifold and reinforces the spike patterns leading to behavioral rewards. The proposed method is evaluated on data collected from two male Sprague Dawley (SD) rats during a two-lever discrimination task. As our contribution is to propose a novel algorithm for neural spike prediction, multiple data segments from two subjects are statistically sufficient to validate the algorithm performance [16], [17], [18]. We train the neural spike prediction models to predict the neural activities from the medial prefrontal cortex (mPFC) to the primary motor cortex (M1), both actively participating in the behavioral task. Results are compared between the models trained with or without the constraint terms. The rest of the paper is organized as follows. Section II introduces the data collection and preprocessing, the proposed constraint method, and the spike prediction model. Section III presents the model training results and compares models with or without the constraint. Section IV concludes this paper.

II. Methods

A. Behavioral Experiments and Data Preprocessing

The experiments presented in this paper were conducted at the Hong Kong University of Science and Technology (HKUST). All animal-handling procedures were approved by the Animal Ethics Committee of HKUST. Two male SD rats, well-trained in a two-lever discrimination task [5], were used for data collection. The task required the rats to make choices and press one of the two levers in response to a random sequence of audio start cues in two different pitches. If the rat pressed and held the correct lever for 500 ms, it would hear a feedback cue in the same pitch as the start cue and be rewarded with water for this successful trial. Pressing the wrong lever or early releasing led to a failed trial with no audio feedback or water reward. The start cue for the coming trial would be given 4-7 seconds after the previous trial ended.

Neural spiking activities from the mPFC and M1 areas of each rat were recorded during the behavioral task by Plexon (Plexon Inc., Dallas, TX). The timing of all the cues and pressing events was recorded by the behavioral recording system (Lafayette Instrument, USA) and synchronized with neural signals. The recorded neural signals were sorted offline to obtain spike trains for single units and binned by 10 ms [5]. To reduce the output dimension of spike prediction models, we selected M1 neurons closely related to the task by finding a minimum set to reconstruct the rats' behavior [13], [19]. All recorded mPFC neurons are used as model inputs to preserve as much as the information related to M1 neural activities and the behavioral task. For Rat A, 17 mPFC neurons were recorded, and nine M1 neurons were selected from 26 recorded neurons. For Rat B, there were 24 recorded mPFC neurons and seven selected M1 neurons out of 12 recorded neurons. For both rats, the recordings contained about 100 successful trials for the low lever and another 100 successful trials for the high lever in an over 40-minute session. These successful trials are preserved for training and testing the transregional prediction model. We randomly shuffle these trials by indexes and split them into five folds to run a five-fold cross-validation. The spike trains of the recorded mPFC neurons serve as the input of spike prediction models to generate spike trains of the selected M1 neurons.

The neural manifold is estimated on a freely moving period, defined as the duration when rats moved around in the behavioral box and did not perform the task. Such a manifold can reflect the background neural firing and is not specific to our designed task. This is critical for disordered subjects with lesions on the neural pathway. Although the subject may fail in behavioral tasks and cannot provide ideal recordings for supervised model training, we can always infer the manifold from these background firings of downstream neurons. Firstly, we smooth the spike trains of the selected M1 neurons with a Gaussian kernel (100 ms standard deviation) to obtain the firing probabilities at each time bin. Then, we extract the latent neural manifold by PCA [20], [21] on the firing probabilities during the freely moving period. This process is a linear projection from the original neural activities to orthogonal, zero-mean principal components (PCs). These PCs define a series of latent variables that reflect statistical features of the neural manifold and are used for designing the constraint terms for the spike generation model based on the reinforcement from behavioral rewards.

B. Behavioral Reinforcement Under the Neural Manifold Constraint

The overall goal of our method is to generate functional neural spike patterns of the M1 neurons from mPFC spike

trains. The generated spike patterns are expected to lead to the desired behavioral movements and fall within the neural manifold. Intuitively, the neural manifold constraint is designed in such a principle that the firing probabilities predicted from the models maintain the statistics on each PC as the targeted neural data. Here, we focus on the first and the second order statistics, i.e., the mean and the variance. Given that the neural manifold estimated from the freely moving period has zero mean and σ_n variance on the n^{th} PC dimension, the RL problem under the neural manifold constraint is formalized as

$$\max \quad J(\theta) \triangleq \sum_k E_{\boldsymbol{y}_k \sim \boldsymbol{p}_k(\theta)} G_k, \tag{1}$$

$$\text{subject to } \mu(l_n) = 0, \tag{2}$$

$$\text{var}(l_n) \le u_n \sigma_n, n = \{1, \ldots, N\}, \tag{3}$$

where N is the number of the M1 neurons, as well as the number of PCs. The objective function $J(\theta)$ is the sum of the expected discounted return, G_k, for the spike trains $\boldsymbol{y}$ following the Bernoulli process defined by firing probabilities $\boldsymbol{p}(\theta)$ [13], where θ represents the weights of the spike generation model (see section C). G_k is defined as $\sum_{t=0}^{D-1} \eta^t R_{k+t}$, where η is the discounted factor, and R_t is the instantaneous reward labeling whether the generated spike trains induce movements for the trial success. The induced movements are estimated by a decoder pretrained from each rat's M1 recordings to the movements. This decoder can be viewed as a proxy for the rat's behavioral response to the micro-stimulations [22], [23], [24]. The $\mu(l_n)$ and $\text{var}(l_n)$ represent the mean and variance over time for the latent variable l_n on the n^{th} PC dimension. For each time bin k, $l_{n,k}$ is obtained by $(\boldsymbol{p}_k - \boldsymbol{m})^T V_n$, where $\boldsymbol{p}_k \in \mathbb{R}^{N\times 1}$ is the predicted instantaneous firing probabilities, $\boldsymbol{m} \in \mathbb{R}^{N\times 1}$ is the mean firing probabilities of the selected M1 neurons, and $V_n \in \mathbb{R}^{N\times 1}$ is the n^{th} column of the PCA projection matrix. Both $\boldsymbol{m}$ and V_n are estimated from the M1 recordings during the freely moving period, defining the coordination and shape of the neural manifold. A relaxation parameter $u_n \ge 1$ is applied to the variance constraint σ_n, considering that neural firing can be more active during the behavioral task than during the freely moving period.

Such an optimization problem with two sets of constraint terms is solved separately. First, the mean squared error (MSE) is a common method for equality constraints in (2). However, in general, MSE is sensitive to outliers. In our experiment, the prediction on spikes can be very noisy and will drift away from the mean values during the task. Thus, we use a log-based measurement and add it into the objective function as a surrogate loss, combining (1) and (2) into

$$\max J(\theta) - \gamma M(\boldsymbol{p}), \tag{4}$$

where $M(\boldsymbol{p}) = \sum_{n=1}^{N} \sum_{k=1}^{K} \left(p_{k,n} \log\left(\frac{p_{k,n}}{m_n}\right) + (1 - p_{k,n}) \log\left(\frac{1-p_{k,n}}{1-m_n}\right)\right)$, $p_{k,n}$ and m_n correspond to the n^{th} neuron of $\boldsymbol{p}_k$ and $\boldsymbol{m}$, respectively, and the hyperparameter γ determines the strength of the constraint.

Second, for inequality constraints on variance values, it is hard to set a constant hyperparameter manually as (4). Thus, we introduce a set of positive Lagrangian multipliers, λ_n, to adapt the constraint strength during training. Then, the optimization problem becomes

$$\min \ -J(\theta) + \gamma M(\boldsymbol{p}) - \sum_{n=1}^{N} \lambda_n (u_n \sigma_n - L_n),$$
$$\text{subject to } \lambda_n \ge 0, n = \{1, \ldots, N\}, \tag{5}$$

where $L_n \triangleq \frac{1}{2K} \sum_{k=1}^{K} (l_{n,k})^2$ is second order statistics for the predicted firing probabilities on the neural manifold.

We solve (5) with the Modified Differential Method of Multipliers (MDMM) [25], which simultaneously performs gradient descent on the model weights and gradient ascent on the Lagrangian multipliers with a damping term controlled by c, shown as

$$\theta \leftarrow \theta + \alpha \Big(\nabla_\theta J(\theta) - \gamma \nabla_\theta M(\boldsymbol{p}) - \sum_{n=1}^{N} \left((\lambda_n - c(u_n \sigma_n - L_n)) \nabla_\theta L_n \right) \Big),$$
$$\lambda_n \leftarrow \max(\lambda_n - \beta(u_n \sigma_n - L_n), 0), \tag{6}$$

where α and β are learning rates. The term $\lambda_n - c(u_n\sigma_n - L_n)$ adaptively balances the learning preference between behavioral performance and neural manifold statistics, where c is a positive constant to reduce the oscillation during training. Intuitively, when generated firing patterns have a large variance on the manifold, $-c(u_n\sigma_n - L_n)$ is positive, giving higher priority to the constraint term $\nabla_\theta L_n$; otherwise, the model will update more towards $\nabla_\theta J(\theta)$ for behavioral performance.

The gradients on the constraint terms are derived as

$$\nabla_\theta M(\boldsymbol{p}) = \frac{1}{K} \sum_{k=1}^{K} \left(\log\left(\frac{p_{k,n}}{m_n}\right) - \log\left(\frac{1-p_{k,n}}{1-m_n}\right)\right) \nabla_\theta \boldsymbol{p}_k,$$
$$\nabla_\theta L_n = \frac{1}{K} \sum_{k=1}^{K} V_n V_n^T (\boldsymbol{p}_k - \boldsymbol{m}) \nabla_\theta \boldsymbol{p}_k, \tag{7}$$

where $\nabla_\theta \boldsymbol{p}_k$ represents the backpropagation for the weights of the spike generation model.

The gradient on the objective function, $\nabla_\theta J(\theta)$, is estimated by the policy gradient [13] as

$$\nabla_\theta J(\theta) = \frac{1}{K} \sum_{k=1}^{K} G_k \nabla_\theta \log\left(P(\boldsymbol{y}_k)\right), \tag{8}$$

where $\log\left(P(\boldsymbol{y}_k)\right)$ is the log-likelihood of spike trains, written as $\boldsymbol{y}_k \log \boldsymbol{p}_k + (1 - \boldsymbol{y}_k) \log(1 - \boldsymbol{p}_k)$. This gradient term can be viewed as a weighted sum of the log-likelihood to guide the model learning more from generated spike trains with higher behavioral rewards.

C. Spike Generation Model

We use a spike-in, spike-out point process model to generate M1 spike trains from the mPFC spike trains. The model has the identical ANN structure as [13] for a fair comparison. Briefly, the model first preprocesses the input spike trains into a history ensemble and then passes it through a 2-layer fully connected ANN. Given N_x input mPFC neurons and H past relevant spikes, the entry in the i^{th} row and j^{th} column of the input ensemble $X_k \in \mathbb{R}^{N_x \times H}$ at time k is expressed as

$$X_k(i, j) = \exp\left(-\frac{k - k_{i,j}}{\tau_j}\right)$$
$$i = 1, \ldots, H; j = 1, \ldots, N_x, \tag{9}$$

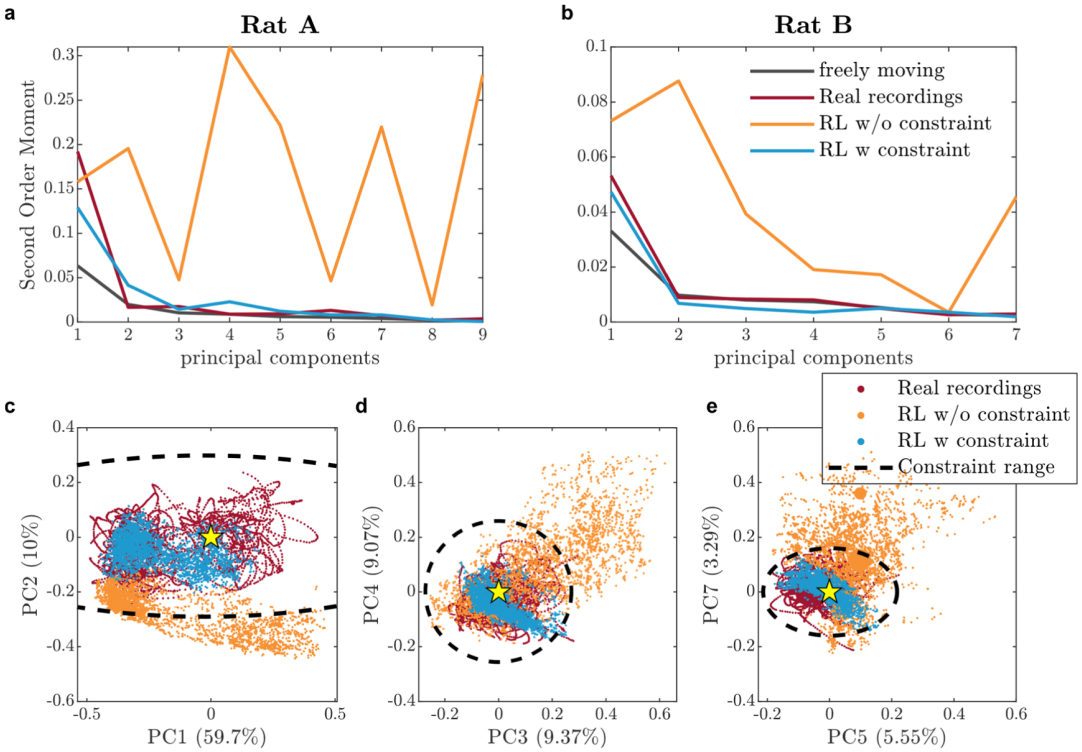

Fig. 1. **Comparing the recorded and the generated neural population activities in the PCA space**. a-b, The second-order moment of the neural activities on the principal components (PCs) for Rat A and Rat B on the test set. The x-axis is the index of PCs, and the y-axis is the second-order moment. The black, red, orange, and blue lines represent the results from the real M1 recordings during the freely moving period, the recordings during the successful trials, the predicted M1 firing probabilities by reinforcement learning (RL) without constraint, and the model predictions by RL with the neural manifold constraint, respectively. c-e, The distribution of Rat B's neural population activity in subspaces spanned by PCs. In each subspace, the two axes correspond to two PCs, and the data points correspond to the firing probabilities of the M1 neural population. The colors of the data points have the same definition as in a-b.

where $k_{i,j}$ is the time index of the i^{th} past spike of the j^{th} input neuron; τ_j is the decay parameter for the j^{th} neuron. Then,

$$\boldsymbol{p}_k = \text{Sig}\left(W_2^{\text{T}}\text{Sig}\left(W_1^{\text{T}} \cdot \text{vec}\{X_k\} + \boldsymbol{b_1}\right) + \boldsymbol{b_2}\right), \tag{10}$$

$$y_{k,n} \sim \text{Bernoulli}(p_{k,n}), \tag{11}$$

where $y_{k,n}$ denotes the generated spike of the n^{th} M1 neuron at time k, and $\theta \triangleq \{W_1, W_2, \boldsymbol{b_1}, \boldsymbol{b_2}\}$ are model weights in compatible sizes. The $\text{Sig}(\cdot)$ is the sigmoid function, and $\text{vec}\{X_k\}$ vectorizes the ensemble X_k for model input.

III. Results

We evaluate the proposed method on the multi-trial neural data from two rats with five-fold cross-validation and compare the model performance with the constraint to those without any constraints. The hyperparameters related to the constraint terms are explored for best performance. We set the learning rate β for Lagrangian multipliers as 1 and the damping constant c as 10. The strength of the entropy regularization γ is explored and set as 0.005, a relatively small value to prevent the predicted firing probabilities from fast convergence to constant values. The relaxation parameters u_n are explored on each rat, following the principle that greater u_n should be given to the first few PCs to allow more space for model exploration. For Rat A, u_1 is set as 3 and u_n for the rest of the PCs is set as 2. For Rat B, u_1 is set as 2 and u_n for the rest is 1, i.e., no relaxation on the other PCs. The other hyperparameters unrelated to the constraint remain the same as [13]. The maximum iteration number is 5000, while the model weights with the best behavioral performance task on the training batch out of all iterations are used for testing. Each model is re-initialized 16 times, and the one with the best performance is used for further analysis.

In the following sections, we will compare the generated neural activity from the trained models with or without the proposed constraint terms. Sections A and B visualize the generated spike train in the feature space and the time domain, respectively. Section C compares the behavioral performance of RL models. Sections D and E discuss the advantages of the neural manifold constraint in the model's robustness and functional neural connectivity, respectively.

A. Neural Manifold Constraint Shapes the Generated Neural Activity in the PCA Space

We first examine the generated neural activities in the PCA space. With the neural manifold constraint, we expect that the RL model predictions have similar distributions to the real recordings in the PCA space, while the RL model without the constraint will have different distributions.

Fig. 1a and b compare the second-order moment of PCs on the test set of Rat A and B, respectively. Here, the black curves show the variances on each PC of the freely moving period M1 recordings. Then, we project the neural population

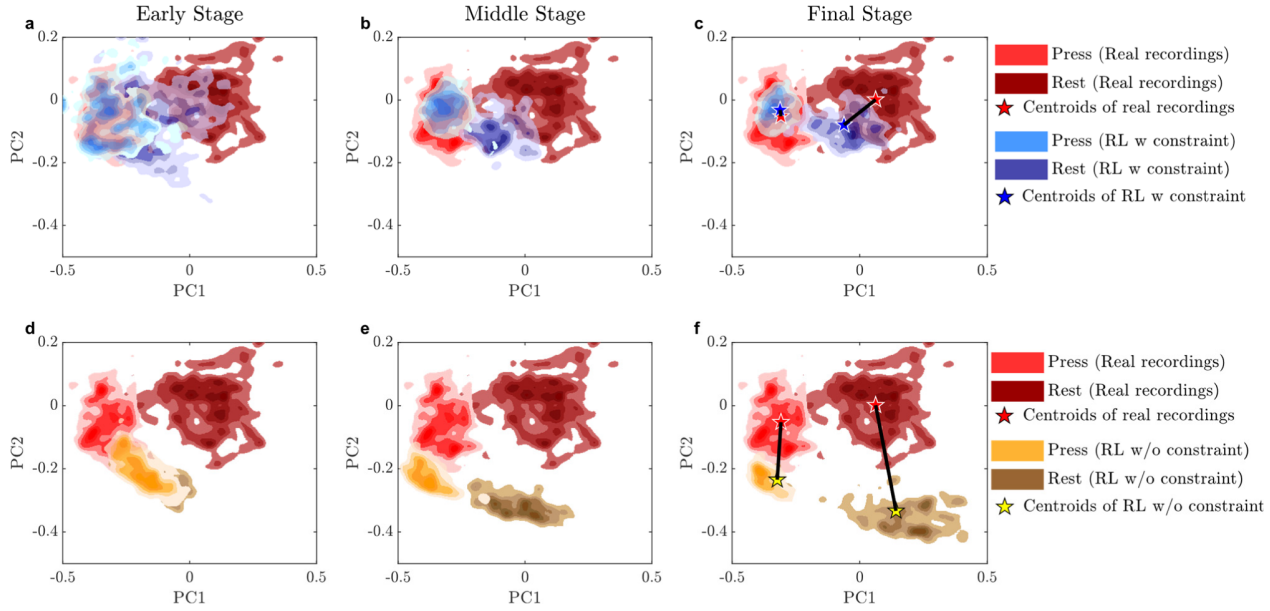

Fig. 2. **The distribution evolvement of neural population activities during model training**. We show the distribution of Rat B's M1 neural population activities in the subspace of the first two PCs. In each subplot, the x-axis and y-axis are the values of PCs. The red and dark red areas correspond to the estimated distributions of recorded M1 neural activities during press and rest movements, respectively. In **a-c**, the light and dark blue areas show the distribution of RL predictions with constraint during press and rest movements. In **d-f**, the yellow and brown areas represent the distribution of RL predictions without constraint during press and rest movements. Specifically, in **c** and **f**, we calculate the centroids of the neural activities during the rest or press movements for different data and label the centroids in corresponding colors. The black lines show the distances between centroids for the same movement.

firing probabilities of real recordings (red), predictions by RL models without constraint (orange), and predictions by RL with constraint (blue) to the same PCA space and show their second-order movements on each PC. The real recordings (black and red curves) show a rapid decrease in the second-order moment, a typical characteristic in neural recordings, indicating that the information is primarily encoded in the low-dimensional space. Note that the real recordings have a larger variance than the freely moving recordings on some PCs, especially the first PC, showing the necessity of relaxation parameters during model training. For RL model predictions, models with the constraint show similar second-order statistics to the real recordings on each PC. However, RL predictions without the constraint have different statistics in the latent space, showing much larger power in most PC dimensions. This suggests that these predictions do not follow the natural patterns of M1 neural firing and may not be suitable for potential micro-stimulations.

We further visualize the M1 firing patterns at each time bin in the PCA space, taking Rat B as an example. In Fig. 1c-e, each subplot shows a subspace spanned by two of the PCs. The numbers in the brackets of axis labels indicate how much variance this PC can be explained in the real recordings of successful trials. The yellow stars show the origins of the subspace, that is, the center of the estimated neural manifold. The dashed ellipses show the range of the neural manifold constraint, whose semi-major and semi-minor axes are defined by $3u_n\sigma_n$ for each PCs. It can be observed that data points from real recordings (red) scatter around the center of the neural manifold. However, the predictions of RL without constraint (orange) are located out of the range of the manifold (black ellipses). By contrast, the RL model with the constraint (blue) generates neural patterns within the neural manifold and close to real recordings (red). In summary, Fig. 1 shows that the RL predictions with the neural manifold constraint have statistics similar to the real recordings. This result indicates that the proposed method successfully shapes the distribution of the RL-generated neural activities.

The emergence of distinct neural patterns can be observed during model training. Fig. 2 shows how the models evolve across different training stages. The three subplots in a row correspond to three training stages of one model, where the subplots a-c and d-f show the evolvement of models with and without the constraint, respectively. In each subplot, we project the neural firing in the test set to the space spanned by the first two PCs, which account for approximately 70% of the total variance of the real recordings. The "press" movement refers to the duration from the lever being pressed to being released, and the "rest" movement refers to the state of rats in 500ms duration before the start cue. The red and dark red areas correspond to the distributions of neural recordings during the press and rest movements, respectively. These areas are estimated by smoothing the data points with a 2-D Gaussian kernel and kept unchanged across subplots, reflecting the neural manifold's coordinate and range. The regions in other colors correspond to the model predictions.

In the early stages of training (61 iterations), neither model exhibits a clear separation in the generated neural activities. This can be seen by the large overlap between light and dark blue areas in Fig. 2a or yellow and brown areas in Fig. 2d. However, the RL predictions with the constraint (Fig. 2a) already reside within the neural manifold (red and

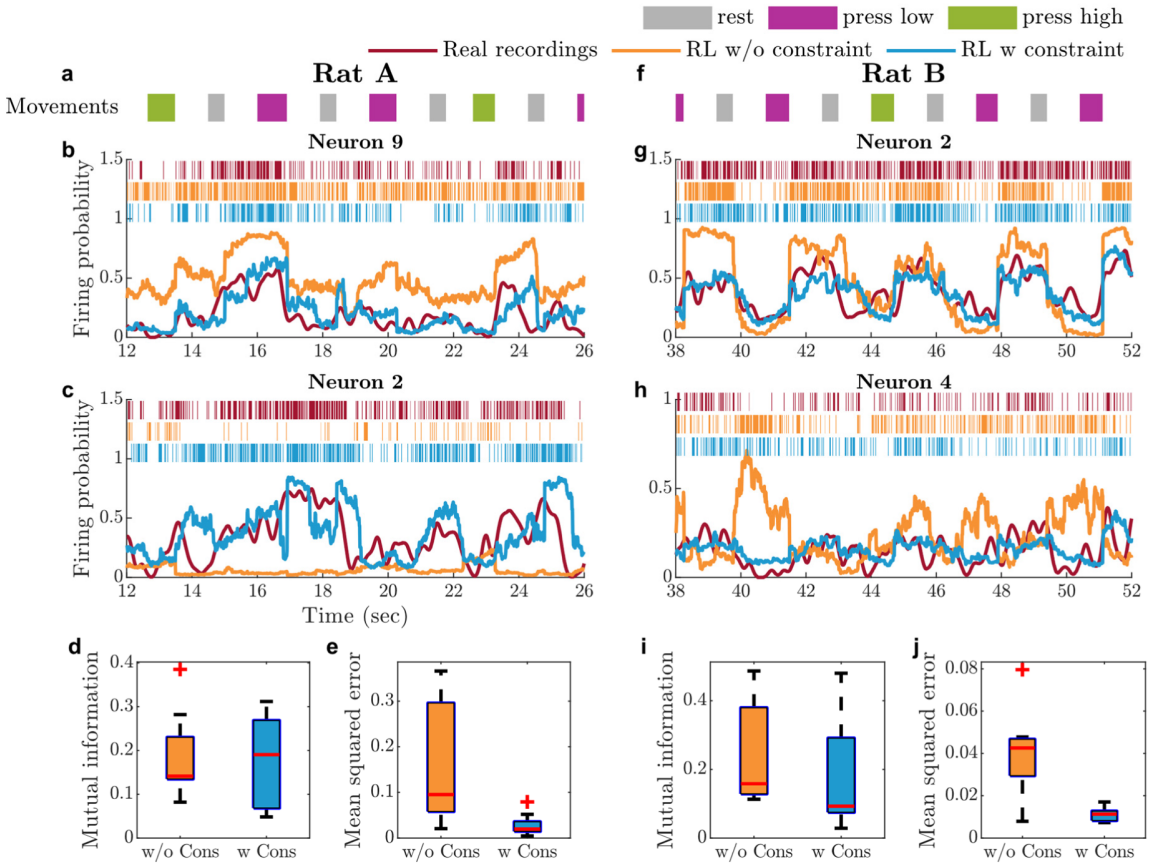

Fig. 3. **Comparing the instantaneous firing patterns between RL models.** **a**, A segment real movements from the test set of Rat A. The x-axis is the time index. The grey, purple, and green bars represent the rest, press low, and press high movements in the task, respectively. **b-c**, Firing probabilities of M1 Neuron 9 and Neuron 2 of Rat A in the same segment as **a**. The y-axis is the firing probability. The vertical bars on the top show the spike trains. The red, orange, and blue colors correspond to the real M1 recordings, RL predictions without the constraint, and RL predictions with the constraint, respectively. **d-e**, Mutual information **(d)** and mean squared error **(e)** between the recorded and the model-predicted firing probabilities. Each box represents the range of the metric values across the nine M1 neurons. The orange boxes correspond to the results between RL predictions without the constraint and the recordings. The blue boxes correspond to the results between the RL predictions with the constraint and the recordings. **f-h**, A segment from the test set of Rat B, showing the behavioral movements and neural activities of M1 Neuron 2 and Neuron 4. **i-j**, Mutual information and mean squared error results across the seven neurons of Rat B.

dark red areas). In contrast, the RL predictions without the constraint (Fig. 2d) already deviate from the real recordings. In the middle training stage (1501 iterations), the RL model with constraint (Fig. 2b) keeps exploring within the neural manifold, gradually learning to generate distinct patterns for different movements. In contrast, the RL model without constraint (Fig. 2e) generates neural activities that significantly deviate from the neural manifold. By the final stage of training, where the models show the best behavioral performance on train sets, the constrained model's predictions (Fig. 2c, 4884 iterations) remain within the neural manifold, while the unconstrained model's predictions (Fig. 2f, 2840 iterations) drift further away. The distances for the centroids of neural activities between real recordings and model predictions are also illustrated by black lines in Fig. 2c and f. The average distance (length of the black lines) for RL predictions with the constraint is 0.08, much better than 0.26 for RL predictions without the constraint. The same metric on Rat A is 0.24 and 0.39 for RL predictions with or without the constraint, where our method also performs better. Overall, the sequential presentation in Fig. 2 provides a clear view of the model's evolvement during training and reveals the role of the neural manifold constraint in generating realistic neural patterns.

B. The Constraint Leads to Natural Instantaneous Firing Patterns

This section compares the instantaneous neural activities generated by RL models with or without the constraint. We first display the behavioral movements of Rat A in Fig. 3a and Rat B in Fig. 3f, which show the rhythmic patterns of rats' behavior. The following subplots illustrate the firing probabilities and spike trains of two neurons from each rat. Results show that the constraint can help RL models generate more natural firing patterns in the time domain.

Firstly, the constraint can reduce the overall power of the neural firing to a more reasonable range. For instance, in Neuron 9 of Rat A (Fig. 3b), the predictions by RL with constraint (blue curve) have smaller amplitude and are closer to real recordings (red curve) than the predictions by RL without constraint (orange curve). This observation is also evident in Rat B's Neuron 2 (Fig. 3g). The predictions of RL

with constraint exhibit lower power during high firing periods (e.g., 38-40s) than that of RL without constraint and maintain the background firing during low firing periods (e.g., 40-42s). These differences make the RL predictions closer to the natural firing patterns.

Secondly, the constraint can shape the modulations of generated neural activity. In Neuron 2 of Rat A (Fig. 3c), the RL model fails to learn the dynamical patterns without the constraint. In contrast, the RL model with the constraint successfully generates modulations similar to real recordings. A more extreme case is Rat B Neuron 4 (Fig. 3h). Here, the neural firing generated by the RL model without the constraint displays an inverse modulation to real recordings, particularly evident in 38-44s and 49-51s. In contrast, the RL model with the constraint learns the correct modulations. This improvement might be because the neural manifold constraint optimizes the model with the correlation information among M1 neural spike trains, thereby enhancing single-neuron prediction.

Then, we statistically evaluate the differences between the recorded and the model-predicted firing probabilities across multiple neurons for each subject. Here, we use the mutual information and MSE to assess the differences between generated and recorded neural activities. Mutual information measures the relative deviation in the distribution of neural firing, while MSE focuses on the absolute differences in temporal patterns. The results are given in Fig. 3d-e for Rat A and i-j for Rat B. For both rats, the mutual information values show no significant difference between the two RL models (two-sided, paired Wilcoxon signed rank test, $p = 0.43$ for Rat A and $p = 0.22$ for Rat B). This means that the firing probabilities predicted by RL with or without constraint do not likely encode information completely different from the recordings, which contributes to explaining why both models can accomplish the task (see Section C). However, when it comes to the MSE between RL predictions and the recordings, the RL models with the constraint are significantly better than the RL models without the constraint (one-sided, paired Wilcoxon signed rank test, $p = 0.002$ for Rat A and $p = 0.016$ for Rat B). On average, the absolute differences are largely reduced by 61.8% for Rat A and 59.7% for Rat B by applying the constraint. These results statistically show that the predicted spike trains are closer to the natural firing in temporal patterns by introducing the neural manifold constraint.

C. Statistical Performance in the Behavioral Tasks

We further examine the statistical behavioral performance of the generated neural spike trains. We decode the predicted M1 spike trains into the three movements, rest, press low, and press high, to see if they can accomplish the behavioral task. Specifically, we expect that the generated spike patterns are decoded into the rest movement before the start cue, and into the press on the correct lever after the cue. The performance is assessed on a time-bin basis and a trial basis, as shown in Fig. 4. The time-bin success rate is the ratio of time bins that the decoded movements match the ground truth, with a chance rate of 0.33 by randomly choosing one of the three

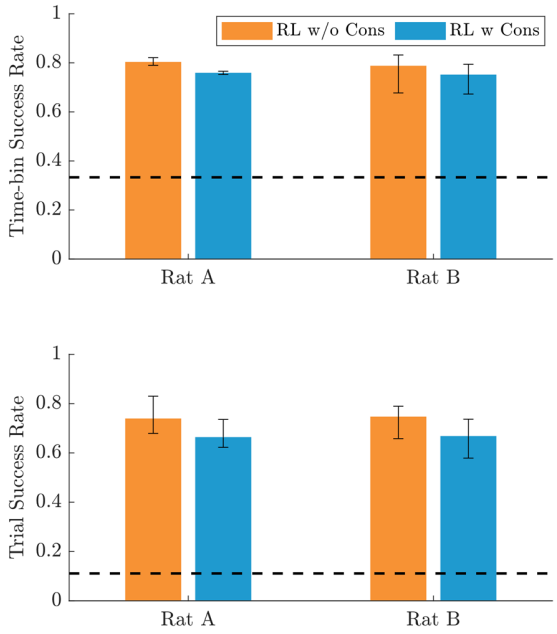

Fig. 4. **Behavior-level performance across five folds**. The upper and lower subplots display the time-bin success rate and the trial success rate of the decoded movements, respectively. Each bar represents the mean and range of the performance on the test sets of five-folds for each rat. The orange and blue bars represent the results from the RL predictions with and without the constraint, respectively, and the dashed lines label the chance rate.

movements. The trial success rate is the ratio of successful trials (correct movement in more than 70% of the time bins within a trial) over all trials in the test set. For trial success, randomly choosing movements in every time bin throughout the trial can barely accomplish the task. Thus, we simplify the trial into a sequence with two steps, randomly selecting one movement before the start cue and another after the cue. This simplified task has a 0.11 chance rate of trial success.

Both RL models significantly beat the chance rate (dashed lines). The time-bin success rates of the RL models are more than 2.26 times the chance rate, and the trial success rates are over 6 times the chance rate. We also notice slight performance drops by introducing the constraint term (about 5% for the time-bin success rate and 10% for the trial success rate). This might be because the RL models need to explore within the neural manifold, instead of generating extreme firing patterns to cater to the given decoders. Overall, the RL model with the constraint can not only generate quite similar neural patterns on the microscopic scale, but also accomplish the task well at the behavioral level.

D. RL Model With Constraint Shows Better Robustness

The neural manifold constraint defines a specific exploration range in the feature space of output patterns. Thus, we anticipate that the RL-generated spike trains under the constraint will converge to similar patterns, regardless of initial states or training data variance. In other words, the RL model with the constraint should be more robust.

Fig. 5a-f present the spike raster and modulations of two M1 neurons, where the first row corresponds to Neuron 4 of

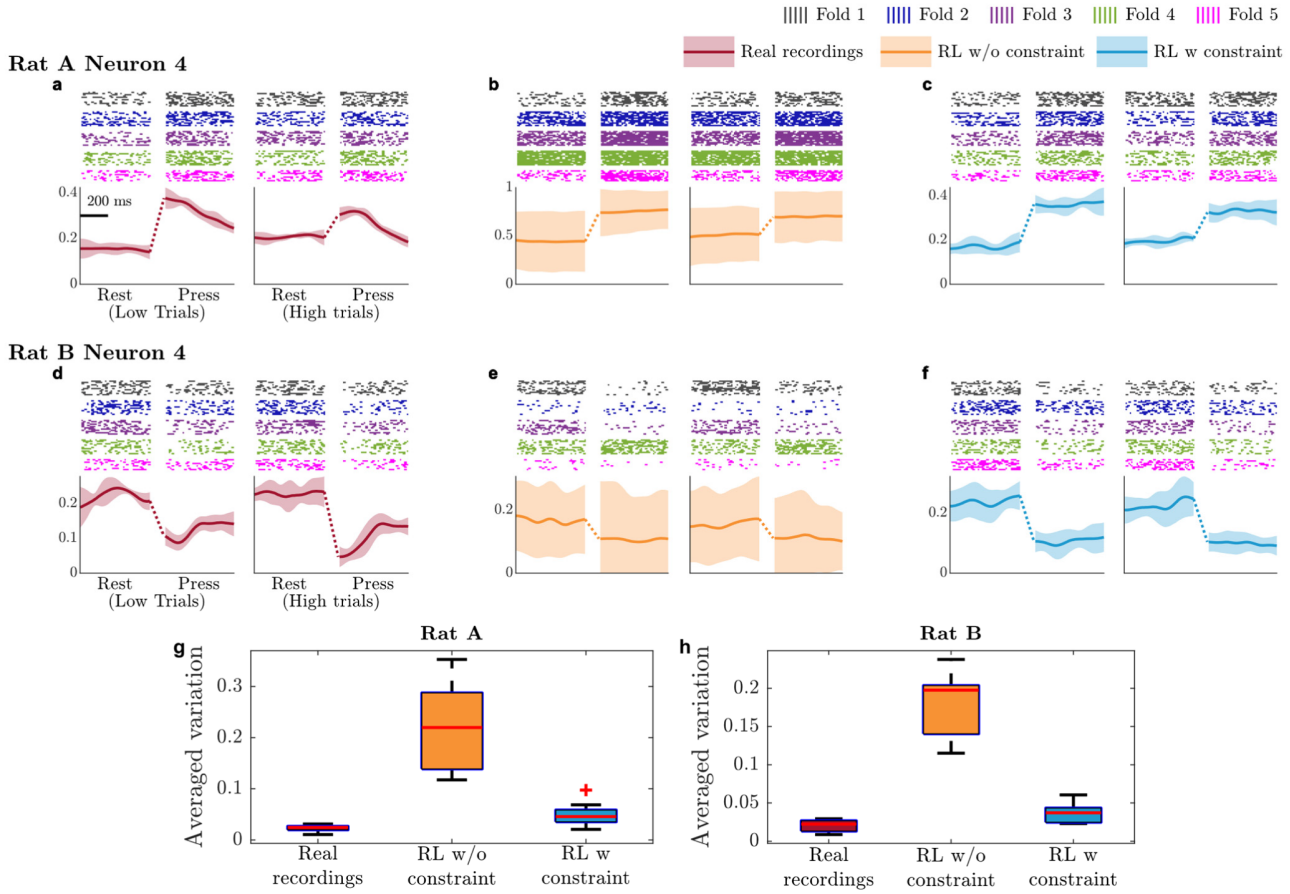

Fig. 5. **Variance of neural modulations across five folds.** **a**-**c**, The modulations of Rat A Neuron 4 for (**a**) real recordings, (**b**) RL predictions without the constraint, and (**c**) RL predictions with the constraint. In each subplot, the left part shows the neural activity in low-lever trials, and the right part shows the high-lever trials. For each part, the x-axis is the time index, where the first 500 ms corresponds to the neural activities before the start cue (rest movement), and the last 500 ms corresponds to the neural activities during the low/high press movement. A gap is added in between the two periods representing the response time of the rats. The y-axis for the raster plots on the top is the trial index, and for the bottom plots is the firing probability. **d**-**f**, The modulations of Rat B Neuron 4 with the same layout as **a**-**c**. **g**-**h**, Statistical comparison for the time-averaged variation of the neural modulations on Rat A and Rat B. The x-axis is the data source, and the y-axis is the averaged variations. Each box shows the results across M1 neurons of one subject. The colors of the boxes have the same definition as a-f.

Rat A, and the second row corresponds to Neuron 4 of Rat B. In each subplot, the raster plots show the spike trains across multiple trials, where the five colors correspond to the five test folds. We then calculate the peri-event firing probability within each fold. The curves with shades under raster plots show the mean $\pm$ standard deviation of the peri-event firing probabilities across folds at each time bin. For real recordings (Fig. 5a and d), the raster plots on top of each subplot display similar firing patterns across all five folds. This results in a small standard deviation of peri-event firing probabilities, as shown by the narrow red shades on the bottom of each subplot. Indeed, the neurons have stationary tuning properties in this well-trained session. However, RL models without constraint (Fig. 5b and e) generate markedly different firing patterns across test folds. For instance, for the raster plots in Fig. 5e, the generated spike trains on folds 1 (black) have higher spike firing during the rest movement than the press movement, i.e., negatively modulated to press. However, the generated spike trains on fold 4 (green) display a positive modulation to press. These differences lead to a significant variance in the neural modulation across folds, as the orange shade shows. In contrast, RL with constraint (Fig. 5c and f) can generate neural firing with similar modulations and low variance, akin to the real recordings.

Fig. 5g-h illustrate the time-averaged variation of the neural modulations by box plots. Such an average value can be regarded as the area of the shades on all movements and normalized by the total time length (2000 ms) in Fig. 5a-f. The recorded spike trains show small variations across folds, representing the fluctuation in real neural firing. Results of RL without constraint (orange) have much larger variations. In contrast, the variations of RL with constraint (blue) are significantly lower than those of RL without constraint, closer to the fluctuation in real recordings (orange). These results suggest that the RL models without the constraint may converge to different policies depending on the training data or initial points. The proposed constraint method substantially reduces the uncertainty in model training, resulting in robust predictions on microscopic neural patterns.

E. The Constraint Induces the Functional Neural Connectivity

Functional neural connectivity is a key feature of neural population firing, which can interpret neural adaptation and

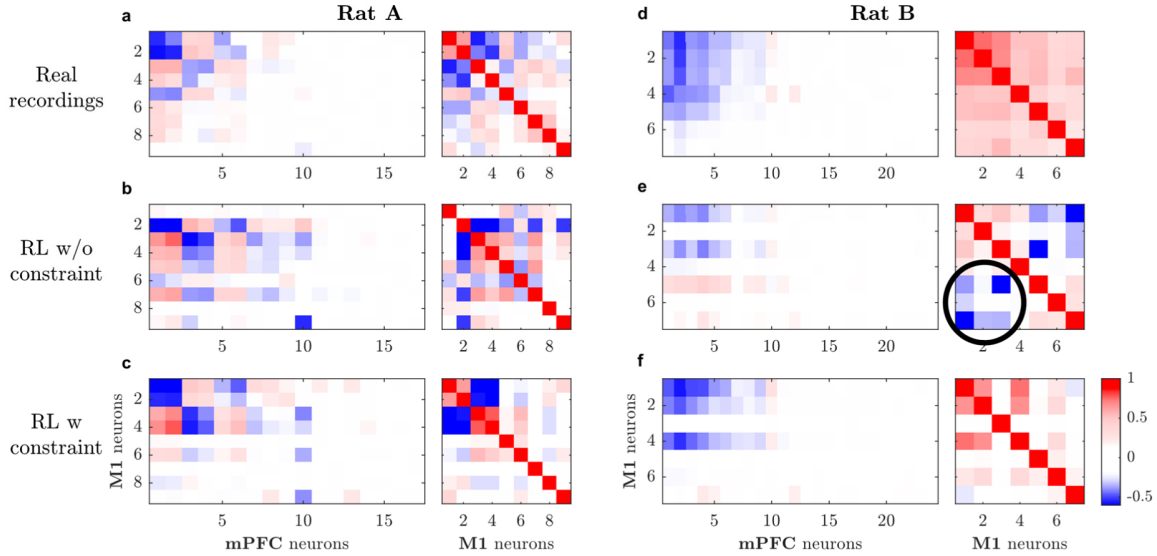

Fig. 6. **Functional correlations among neurons**. a, b, and c show the correlation matrix estimated from Rat A's M1 spike trains of the real recordings, RL predictions without constraint, and RL predictions with constraint, respectively. Similarly, d, e, and f show the estimated correlation matrix of Rat B. In each subplot, the x-axis for the left/right image is the index of mPFC/M1 neurons, and the y-axis is the index of M1. The indexes of the neurons are rearranged in the descending order of the sum of absolute CC values in real recordings.

improve neural decoding in BMI systems [26], [27]. The neural manifold is regarded as a statical representation of neural correlations, which results from the connectivity among cortical networks. Therefore, we hypothesize that the neural manifold constraint can induce functional neural connectivity within the RL-generated neural activities.

We estimate the functional connectivity with the correlation coefficient (CC) between the firing probabilities of two neurons. Fig. 6 shows these pairwise correlations among mPFC and M1 neural firing probabilities for Rat A (left column) and Rat B (right column). The CC values are computed on each test fold and averaged across five folds as the weights of the functional connectivity among neurons. The firing probabilities of mPFC neurons are smoothed from the recorded spike trains, and the firing probabilities of M1 neurons are either smoothed from the recordings (Fig. 6a, d) or generated by the RL models. Within each subplot, the rectangle on the left shows the correlations between mPFC and M1 neural activities, and the square on the right shows the inner correlations among M1 neural activities. The red color represents positive correlations, and the blue represents negative correlations, where a darker color means a larger CC value and indicates a higher functional connectivity weight.

We notice that the correlation matrix for the predictions with constraint (Fig. 6c, f) resembles the correlation among real recordings (Fig. 6a, d), while the predictions without constraint (Fig. 6b, e) show different correlations. For instance, the square matrix in Fig. 6d is mainly filled in red, showing that all selected M1 neurons of Rat B positively correlate with each other in recordings. In contrast, the generated neural activities by RL models without constraint have negative correlations among neurons, shown by the blue blocks inside the black circle in Fig. 6e. After applying the constraint (Fig. 6f), the correlations among M1 neurons are dominated by positive again. Overall, the mean squared errors (MSE) between the correlation matrices of real recordings and the one of RL with constraint are 0.061 and 0.039 for Rat A and Rat B, respectively, which is much better than the MSE for RL without constraint (0.117 and 0.099 for Rat A and Rat B, respectively). These results indicate that the proposed constraint method can induce realistic neuronal correlations.

IV. Conclusion

Bi-directional neural prostheses restore the damaged neural pathway through electrical stimulations on downstream neurons. Spike prediction models that predict downstream neural spiking activities from upstream are essential for generating spatiotemporal stimulation patterns. When ground truth neural firing is unavailable due to neural pathway damages, the models can be trained using RL methods. However, previous studies neglect the intrinsic correlations among neurons and the constraints in neural firing patterns, thus may result in firing patterns deviating from the natural neural activities. Using these patterns as stimulations may hinder the subject's learning and cause safety concerns. In this paper, we propose the neural manifold constraint for RL-based spike prediction models and evaluate our method on the spike prediction from mPFC to M1 neurons of SD rats during the two-lever discrimination task. Results show that these novel constraint terms successfully shape the generated neural activity into patterns similar to real recordings. This resemblance is found in both neural population activities and individual neurons' firing probabilities. Compared with previous methods, the neural manifold constraint reduces the mean squared error between the predicted firing probabilities and the real recordings by 61%. Besides, the models with the constraint reach similar behavioral performance as those without the constraints, but become more robust in generated firing patterns across multiple data segments. Moreover, since the neural manifold is a statistical representation of the connectivity among neural populations, our constraint method ultimately induces realistic correlations among output neurons and between input and

output neurons. In summary, the proposed RL models with the neural manifold constraint simultaneously achieve high behavioral performance and better microscopic neural patterns. Our method is an important step toward the clinical usage of RL-based spike prediction models for generating neural stimulation patterns.

Future research could be conducted in two aspects. Firstly, current constraint terms only describe the spatial correlations among neurons and are evaluated on discrete movements. However, the temporal dynamics in neural population activities are also critical for transregional communication functions, especially when precisely controlling the complex limb trajectories in the three-dimensional space [28], [29]. Our models should capture these neural dynamics for continuous behavior in real-world applications. Secondly, long-term learning may change the neural manifold over time [30]. It is essential to adapt to these non-stationary changes. For instance, in the simplest case, the estimated neural manifold needs periodical recalibration to update the constraint terms. These approaches will lead to dynamic and adaptive spike prediction models for neural prostheses.

REFERENCES

[1] A. N. Belkacem, N. Jamil, S. Khalid, and F. Alnajjar, "On closed-loop brain stimulation systems for improving the quality of life of patients with neurological disorders," *Frontiers Human Neurosci.*, vol. 17, pp. 1–23, Mar. 2023.

[2] A. B. Ajiboye et al., "Restoration of reaching and grasping movements through brain-controlled muscle stimulation in a person with tetraplegia: A proof-of-concept demonstration," *Lancet*, vol. 389, no. 10081, pp. 1821–1830, May 2017.

[3] M. Capogrosso et al., "A brain-spine interface alleviating gait deficits after spinal cord injury in primates," *Nature*, vol. 539, pp. 284–288, Aug. 7628.

[4] R. E. Hampson et al., "Developing a hippocampal neural prosthetic to facilitate human memory encoding and recall," *J. Neural Eng.*, vol. 15, no. 3, Jun. 2018, Art. no. 036014.

[5] S. Wu et al., "Spike prediction on primary motor cortex from medial prefrontal cortex during task learning," *J. Neural Eng.*, vol. 19, no. 4, Aug. 2022, Art. no. 046025.

[6] C. Qian, X. Sun, Y. Wang, X. Zheng, Y. Wang, and G. Pan, "Binless kernel machine: Modeling spike train transformation for cognitive neural prostheses," *Neural Comput.*, vol. 32, no. 10, pp. 1863–1900, Oct. 2020.

[7] S. Wu and Y. Wang, "Applying neural manifold constraint on point process model for neural spike prediction," in *Proc. 45th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Jun. 2023, pp. 1–4.

[8] J. P. Cunningham and B. M. Yu, "Dimensionality reduction for large-scale neural recordings," *Nature Neurosci.*, vol. 17, no. 11, pp. 1500–1509, Nov. 2014.

[9] P. T. Sadtler et al., "Neural constraints on learning," *Nature*, vol. 512, no. 7515, pp. 423–426, 7515.

[10] E. Wärnberg and A. Kumar, "Perturbing low dimensional activity manifolds in spiking neuronal networks," *PLOS Comput. Biol.*, vol. 15, no. 5, May 2019, Art. no. e1007074.

[11] J. A. Gallego, M. G. Perich, R. H. Chowdhury, S. A. Solla, and L. E. Miller, "Long-term stability of cortical population dynamics underlying consistent behavior," *Nature Neurosci.*, vol. 23, no. 2, pp. 260–270, Feb. 2020.

[12] J. A. Gallego, M. G. Perich, L. E. Miller, and S. A. Solla, "Neural manifolds for the control of movement," *Neuron*, vol. 94, no. 5, pp. 978–984, Jun. 2017.

[13] S. Wu et al., "Generative neural spike prediction from upstream neural activity via behavioral reinforcement," *BioRxiv*, 2023, doi: 10.1101/2023.07.25.550495. [Online]. Available: https://www.biorxiv.org/content/early/2023/07/28/2023.07.25.550495

[14] D. Sussillo, M. M. Churchland, M. T. Kaufman, and K. V. Shenoy, "A neural network that finds a naturalistic solution for the production of muscle activity," *Nature Neurosci.*, vol. 18, no. 7, pp. 1025–1033, Jul. 2015.

[15] J. A. Michaels, S. Schaffelhofer, A. Agudelo-Toro, and H. Scherberger, "A goal-driven modular neural network predicts parietofrontal neural dynamics during grasping," *Proc. Nat. Acad. Sci. USA*, vol. 117, no. 50, pp. 32124–32135, Dec. 2020.

[16] A. Ergun, R. Barbieri, U. T. Eden, M. A. Wilson, and E. N. Brown, "Construction of point process adaptive filter algorithms for neural systems using sequential Monte Carlo methods," *IEEE Trans. Biomed. Eng.*, vol. 54, no. 3, pp. 419–428, Mar. 2007.

[17] R. Barbieri et al., "Dynamic analyses of information encoding in neural ensembles," *Neural Comput.*, vol. 16, no. 2, pp. 277–307, Feb. 2004.

[18] E. N. Brown, D. P. Nguyen, L. M. Frank, M. A. Wilson, and V. Solo, "An analysis of neural receptive field plasticity by point process adaptive filtering," *Proc. Nat. Acad. Sci. USA*, vol. 98, no. 21, pp. 12261–12266, Oct. 2001.

[19] Y. Wang, J. C. Principe, and J. C. Sanchez, "Ascertaining neuron importance by information theoretical analysis in motor brain–machine interfaces," *Neural Netw.*, vol. 22, nos. 5–6, pp. 781–790, Jul. 2009.

[20] J. A. Gallego, M. G. Perich, S. N. Naufel, C. Ethier, S. A. Solla, and L. E. Miller, "Cortical population activity within a preserved neural manifold underlies multiple motor behaviors," *Nature Commun.*, vol. 9, no. 1, pp. 13–4233, Oct. 2018.

[21] S. Zheng et al., "A spiking neural network based on neural manifold for augmenting intracortical brain-computer interface data," in *Proc. Int. Conf. Artif. Neural Netw.*, 2022, pp. 519–530.

[22] J. S. Choi, A. J. Brockmeier, D. B. McNiel, L. M. V. Kraus, J. C. Príncipe, and J. T. Francis, "Eliciting naturalistic cortical responses with a sensory prosthesis via optimized microstimulation," *J. Neural Eng.*, vol. 13, no. 5, Oct. 2016, Art. no. 056007.

[23] M. J. Bryan, L. P. Jiang, and R. P. N. Rao, "Neural co-processors for restoring brain function: Results from a cortical model of grasping," *J. Neural Eng.*, vol. 20, no. 3, Jun. 2023, Art. no. 036004.

[24] S. S. Nagrale, A. Yousefi, T. I. Netoff, and A. S. Widge, "In silico development and validation of Bayesian methods for optimizing deep brain stimulation to enhance cognitive control," *J. Neural Eng.*, vol. 20, no. 3, Jun. 2023, Art. no. 036015.

[25] J. Platt and A. Barr, "Constrained differential optimization," in *Proc. Adv. Neural Inf. Process. Syst.*, 1987, pp. 1–26.

[26] M. Li, S. Chen, X. Liu, Z. Song, and Y. Wang, "Modeling neural connectivity in a point-process analogue of Kalman filter," in *Proc. 44th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2022, pp. 768–771.

[27] S. Chen, X. Liu, and Y. Wang, "Considering neural connectivity in point process decoder for brain-machine interface," in *Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Nov. 2021, pp. 6341–6344.

[28] H. Abbaspourazad, M. Choudhury, Y. T. Wong, B. Pesaran, and M. M. Shanechi, "Multiscale low-dimensional motor cortical state dynamics predict naturalistic reach-and-grasp behavior," *Nature Commun.*, vol. 12, no. 1, p. 607, Jan. 2021.

[29] H. Abbaspourazad, E. Erturk, B. Pesaran, and M. M. Shanechi, "Dynamical flexible inference of nonlinear latent factors and structures in neural population activity," *Nature Biomed. Eng.*, vol. 8, no. 1, pp. 85–108, Dec. 2023.

[30] E. R. Oby et al., "New neural activity patterns emerge with long-term learning," *Proc. Nat. Acad. Sci. USA*, vol. 116, no. 30, pp. 15210–15215, Jul. 2019.