31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

A Regularized Point Process Generalized Linear Model for Assessing the Functional Connectivity in the Cat Motor Cortex

Zhe Chen, David F. Putrino, Demba E. Ba, Soumya Ghosh, Riccardo Barbieri, and Emery N. Brown

Abstract— **Identification of multiple simultaneously recorded neural spike train recordings is an important task in understanding neuronal dependency, functional connectivity, and temporal causality in neural systems. An assessment of the functional connectivity in a group of ensemble cells was performed using a regularized point process generalized linear model (GLM) that incorporates temporal smoothness or contiguity of the solution. An efficient convex optimization algorithm was then developed for the regularized solution. The point process model was applied to an ensemble of neurons recorded from the cat motor cortex during a skilled reaching task. The implications of this analysis to the coding of skilled movement in primary motor cortex is discussed.**

I. Introduction

Identifying a neuronal system via multivariate neural spike trains recorded from ensemble neurons has many valuable implications for understanding the system from a statistical perspective, and has been used for establishing statistical associations or causality between neurons, or finding spatiotemporal correlations, or studying the functional connectivity [3], [8], [9], [10], [12], [14]. A statistical treatment of multiple spike trains is to use the theory of stochastic multivariate point processes. Statistical inference for point process observations often starts with a certain class of statistical model, followed by parameter estimation by either maximum likelihood or Bayesian inference procedure [1], [2], [8], [14], [16].

The point process generalized linear model (GLM) [10], [16] has been widely used for characterizing functional (spiking) dependence among ensemble neurons. Recent new developments allow for the incorporation of the Bayesian inference [14] or modeling common input [8]. Here we propose a regularized point process GLM that imposes a temporal smoothness constraint on the parameter space to explore a physiologically plausible solution. The regularized solution can be casted within a convex optimization framework and can be solved efficiently by a linear conjugate gradient method. The regularized point process GLM was applied to real ensemble neurons recorded from awake behaving cats during a reaching task [13]. The functional connectivity and spiking temporal dependence regarding different stages and different task performances were examined, and some physiological interpretations and discussions on the new findings of this pilot study were presented.

II. A Point Process Model for Multiple Spike Trains

Let $c = 1, \cdots, C$ denote the index of a multivariate (C-dimensional) point process. For the cth point process, let $N_c(t)$ denote the counting process up to time t, and let $\mathrm{d}N_c(t)$ denote the indicator variable, which equals to 1 if there is an event (spike) at time t and 0 otherwise. Therefore, the multiple neural spike train data are completely characterized by a multivariate point process $N_{1:C}(0:T)$.

This work was supported by NIH Grants DP1-OD003646, R01-DA015644 and R01-HL084502.

The authors are with the Neuroscience Statistics Research Laboratory, Massachusetts General Hospital, Harvard Medical School, Boston, MA 02114, USA. S. Ghosh is with the School of Medicine and Pharmacology, Perth, University of Western Australia, Australia. (Email: zhechen@mit.edu)

In modeling the neural spike train point process, the *conditional intensity function* (CIF) is used to characterize the instantaneous firing probability of a discrete event (i.e., spike) [1]:

$$\lambda_c(t|\mathcal{H}_{0:t}) = \lim_{\Delta \to 0} \frac{\Pr\{N_c(t+\Delta) - N_c(t) = 1|\mathcal{H}_{0:t}\}}{\Delta}, \qquad (1)$$

where $\mathcal{H}_{0:t}$ denotes all of ensemble neuronal firing history and any other information up to time t. Where Δ is sufficiently small, the product CIF $\lambda_c(t|\mathcal{H}_{0:t})\Delta$ tells approximately the probability of observing a spike within the interval $[t, t+\Delta)$:

$$\Pr\{N_c(t+\Delta) - N_c(t) = 1|\mathcal{H}_{0:t}\} \approx \lambda_c(t|\mathcal{H}_{0:t})\Delta. \qquad (2)$$

Here, we restrict ourselves to the cases where $\sum_{c=1}^{C} \mathrm{d}N_c(t) \leq 1$ at any time t, i.e., no joint firing is allowed in the continuous-time setting (in the case of discrete-time setting, no joint firing is allowed at the finest temporal scale under consideration). Let $\boldsymbol{\theta}$ denote the ensemble unknown parameters in the parametric form of function $\{\lambda_c\}_{c=1}^{C}$. Specifically, we express the CIF in the following log-linear form [10]:[1]

$$\log \lambda_c(t) = \boldsymbol{\alpha}\boldsymbol{x}^c(t) = \sum_{j=0}^{d} \alpha_j x_j^c(t) = \alpha_{i,0} + \sum_{i=1}^{C}\sum_{k=1}^{K} \alpha_{i,k} x_{i,t-k}^c$$

where $dim(\boldsymbol{\alpha}) = d+1$ (where $d = C \times K$) denotes total number of parameters in the augmented parameter vector $\boldsymbol{\alpha} = \{\alpha_{i,k}\}$, and $\boldsymbol{x}^c(t) = \{x_{i,t-k}^c\}$, where $x_{i,0}^c = 1$ and $x_{i,t-k}^c$ denotes the spike count information from cell i at the kth time-lag history windows.

Let $\boldsymbol{\theta} = \{\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_C\}$, where $dim(\boldsymbol{\theta}) = C(1+d)$. By assuming that the spike trains are mutually *conditionally* independent, the continuous-time log-likelihood of observed data is given by:

$$L(\boldsymbol{\theta}) = \sum_{c=1}^{C} \left\{ \int_0^T -\lambda_c(t|\boldsymbol{\theta})\mathrm{d}t + \int_0^T \log(\lambda_c(t|\boldsymbol{\theta}))\mathrm{d}N_c(t) \right\}. \qquad (3)$$

By discretization of (3), we also obtain the discrete-time log-likelihood function, in which the integration will be replaced by a finite sum. From (3) it is clear that $-L(\boldsymbol{\theta})$ is convex with respect to (w.r.t.) to each λ_c. In addition, the index c is uncoupled from each other in the network log-likelihood function, which implies that we can optimize the function separately for each spike train $N_c(0:T)$ once $\lambda_c(t)$ is specified. For simplicity, from now on we will drop off the index c at λ_c and $\boldsymbol{\alpha}_c$ when no confusion occurs.

III. Regularization and Optimization

When the size of parameter space is large, it is often desirable to impose certain prior knowledge (such as spatial sparsity) or physiologically plausible constraint (such as temporal smoothness) on the parameters [15], [6]. This can be done by the so-called "regularization" to improve the generalization ability of the model (on unseen data) while fitting finite training data. Regularization can be interpreted as imposing a prior on the parameter space in terms of Bayesian inference, and the log-likelihood will be interpreted as the

[1] Note that here we use the simplified notation: $\lambda_c(t) = \lambda_c(t|\boldsymbol{\alpha}, \mathcal{H}_{0:t})$.

log posterior density of the parameters [14]. Regularization seeks to maximize a penalized log-likelihood function, which consists of a log-likelihood function plus a penalty function weighted by a regularization parameter. Specifically, we propose to maximize the following penalized log-likelihood function using ℓ_2 regularization:

$$L(\boldsymbol{\alpha}) = \int_0^T -\lambda(t|\boldsymbol{\alpha})\mathrm{d}t + \int_0^T \log(\lambda(t|\boldsymbol{\alpha}))\mathrm{d}N(t) - \rho\boldsymbol{\alpha}^T\boldsymbol{Q}\boldsymbol{\alpha} \quad (4)$$

where $\rho > 0$ denotes a regularization parameter, and $\boldsymbol{Q}$ denotes a user-defined positive semi-definite matrix (to be defined below). When $\boldsymbol{Q} = \mathbf{I}$ (identity matrix), then the standard "ridge regression" is recovered. Maximization of the penalized log-likelihood (4) can be solved by the expectation-maximization (EM) algorithm [5], below we propose an alternative yet more efficient solution.

A standard way to optimize the penalized log-liklehood is through the Newton method. Specifically, let $\mathbf{H}(\boldsymbol{\alpha})$ and $\mathbf{g}(\boldsymbol{\alpha})$ denote the Hessian matrix and gradient vector of the parameter vector $\boldsymbol{\alpha}$, respectively, computed from the $L(\boldsymbol{\alpha})$ given by (4); the iterative Newton update equation (at the nth step) is given by

$$\begin{aligned}\boldsymbol{\alpha}_{n+1} &= \boldsymbol{\alpha}_n - \mathbf{H}^{-1}(\boldsymbol{\alpha}_n)\mathbf{g}(\boldsymbol{\alpha}_n) \\ &= \boldsymbol{\alpha}_n + \left[\boldsymbol{X}^T\boldsymbol{W}(\boldsymbol{\alpha}_n)\boldsymbol{X} + \rho\boldsymbol{Q}\right]^{-1}\boldsymbol{X}^T(\boldsymbol{y} - \hat{\boldsymbol{y}}(\boldsymbol{\alpha}_n)) \quad (5)\end{aligned}$$

where $\boldsymbol{y} = [\mathrm{d}N(1), \ldots, \mathrm{d}N(T)]$, $\boldsymbol{X} = [\boldsymbol{x}(1), \ldots \boldsymbol{x}(T)]$, $\boldsymbol{W}(\boldsymbol{\alpha}) = diag\{w_1, \ldots, w_T\}$ is a $T \times T$ diagonal weighting matrix ($w_t = \lambda(t; \boldsymbol{\alpha})$). Since $L(\boldsymbol{\alpha})$ is concave w.r.t. $\boldsymbol{\alpha}$, the maximum likelihood estimation reduces to a convex optimization problem. Equation (5) can also be formulated as iteratively solving a linear quadratic system: $\left[\boldsymbol{X}^T\boldsymbol{W}(\boldsymbol{\alpha}_n)\boldsymbol{X} + \rho\boldsymbol{Q}\right]\boldsymbol{\alpha}_{n+1} = \boldsymbol{X}^T\boldsymbol{W}(\boldsymbol{\alpha}_n)\boldsymbol{b}$, where $\boldsymbol{b} = \boldsymbol{X}\boldsymbol{\alpha}_n + \boldsymbol{W}^{-1}(\boldsymbol{\alpha}_n)(\boldsymbol{y} - \hat{\boldsymbol{y}}(\boldsymbol{\alpha}_n))$. For such a convex optimization problem, efficient iterative algorithms such as *reweighted least squares* (RWLS) [11] or conjugate gradient [7] can be applied.

Next, for the purpose of regularization we want to impose a physiologically-inspired constraint on the parameters in $\boldsymbol{\alpha}$. The motivation is that spiking history dependence often has a local smoothness between the neighboring temporal windows. When the parameter sequences $\{\alpha_{c,k}\}$ are temporally smooth for any index c, the local variance will be relatively small. Let $\overline{\alpha}_{c,k}$ denote the corresponding short-term exponentially weighted average of $\alpha_{c,k}$, namely

$$\overline{\alpha}_{c,k} = \gamma\overline{\alpha}_{c,k-1} + (1-\gamma)\alpha_{c,k}$$

where $0 < \gamma < 1$ is a forgetting factor that determines the range of local smoothness. Let us define a new quadratic penalty function:

$$\sum_c \sum_k (\alpha_{c,k} - \overline{\alpha}_{c,k})^2 = \sum_c \sum_k \gamma(\alpha_{c,k} - \overline{\alpha}_{c,k-1})^2, \quad (6)$$

which penalizes the local variance of $\{\alpha_{c,k}\}$. Let $\overline{\boldsymbol{\alpha}}_c$ denote the short-term average vector for the corresponding parameter $\boldsymbol{\alpha}_c = [\alpha_{c,1}, \ldots, \alpha_{c,K}]$, then we further have

$$\|\boldsymbol{\alpha}_c - \overline{\boldsymbol{\alpha}}_c\|^2 = \|\boldsymbol{\alpha}_c - \mathbf{S}\boldsymbol{\alpha}_c\|^2, \quad (7)$$

where a smoothing matrix $\mathbf{S}$ is introduced to represent $\overline{\boldsymbol{\alpha}}_c$ in terms of $\boldsymbol{\alpha}_c$. Note that the exponentially moving average $\overline{\alpha}_{c,k}$ can be viewed as a *convolution product* between the sequences $\{\alpha_{c,k}\}$ and a template. Suppose the template vector has an exponential-decay property with length ℓ=4, such that $template = [(1-\gamma), \gamma(1-\gamma), \gamma^2(1-\gamma), \gamma^3(1-\gamma)]$. Note that the convolution smoothing operation can also be conveniently expressed as a matrix product

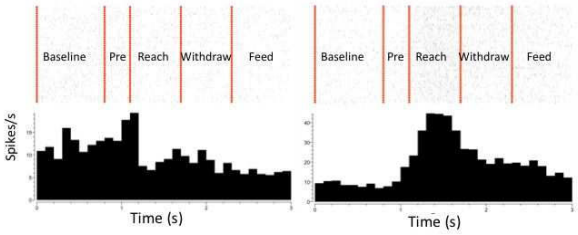

Fig. 1. Spike rasters from all 42 trials and the peri-stimulus time histograms (PSTHs, binwidth 100 ms) from two task-related M1 pyramidal neurons (#6 and 7). Red lines mark the average timing of different stages of the task.

operation: $\overline{\boldsymbol{\alpha}}_c = \mathbf{S}\boldsymbol{\alpha}_c$, where $\mathbf{S}$ is a Toeplitz matrix with the right-shifted template appearing at each row given as follows:

$$\mathbf{S} = \begin{bmatrix} 1-\gamma & 0 & 0 & 0 & \cdots & 0 \\ \gamma(1-\gamma) & 1-\gamma & 0 & 0 & \cdots & 0 \\ \gamma^2(1-\gamma) & \gamma(1-\gamma) & 1-\gamma & 0 & \cdots & 0 \\ \gamma^3(1-\gamma) & \gamma^2(1-\gamma) & \gamma(1-\gamma) & 1-\gamma & 0 & \cdots \\ 0 & \gamma^3(1-\gamma) & \gamma^2(1-\gamma) & \gamma(1-\gamma) & 1-\gamma & \cdots \\ & & \cdots & 0 & \ddots & \vdots \end{bmatrix}$$

Finally, we obtain the regularization matrix $\boldsymbol{Q} = \mathbf{P}^T\mathbf{P}$, where in light of (7) $\mathbf{P}$ has a block-Toeplitz structure:

$$\mathbf{P} = \begin{bmatrix} \mathbf{I}-\mathbf{S} & 0 & \cdots & 0 \\ 0 & \mathbf{I}-\mathbf{S} & \cdots & 0 \\ & & \ddots & \vdots \\ 0 & \cdots & \ddots & \vdots \\ 0 & \cdots & 0 & \mathbf{I}-\mathbf{S} \end{bmatrix}. \quad (8)$$

Our smoothing operator can be seen as an extension of the contingent smoothness operator (as in [15]), where the term $(\alpha_{c,k} - \overline{\alpha}_{c,k})^2$ in (8) is replaced by $(\alpha_{c,k} - \alpha_{c,k-1})^2$ (i.e., the local mean $\overline{\alpha}_{c,k}$ is replaced by its intermediate neighbor $\alpha_{c,k-1}$ without moving averaging). Nevertheless, our "smoothness" operator is more general and also accommodates [6] as a special case. Like ours, the regularization matrix $\boldsymbol{Q}$ in [6] also has a block-Toeplitz structure. Since the regularized minimization problem is convex w.r.t. $\boldsymbol{\alpha}$, so the final regularized solution remains globally optimal.

Upon convergence of the algorithm (using a criterion that the log-likelihood change in two subsequent updates is less than 10^{-3}), the goodness-of-fit of the point process model is evaluated based on the *Time-Rescaling Theorem* and *Kolmogorov-Smirnov* (KS) test [1]. The autocorrelation function of the rescaled time series is also computed to test the independence. It should be pointed out that our optimization procedure is insensitive to the initial parameter vector $\boldsymbol{\alpha}$, and it typically converges within 10 iterations. In our experiment, we empirically set γ=0.5, and the regularization parameter ρ can be selected by cross-validation on the KS statistic or the deviance.

IV. Data and Results

The real spike train data used in this study were recorded from 13 neurons in forelimb (7/13) and hindlimb (6/13) areas of the primary motor cortex (M1) of awake behaving cats in an extracellular recording preparation [13]. During recording periods, the animals performed a skilled reaching task. The data analyzed in this study were taken from one recording session in one cat where the same task was repeated for 42 independent trials. Each complete trial recording lasted 3 seconds, which contains "baseline", "pre-movement", "reaching", "withdraw", and "feed" [13]. Thirteen single units were isolated, including 8 regular-spiking (RS, pyramidal) cells and 5 fast-spiking (FS, interneuron) cells (# 3, 4, 5, 8, 10), which were classified based on baseline firing rate and spike wavelength [4]. Based on whether the cat was able to accurately grasp the food pellet in one, or required more than one attempt, the

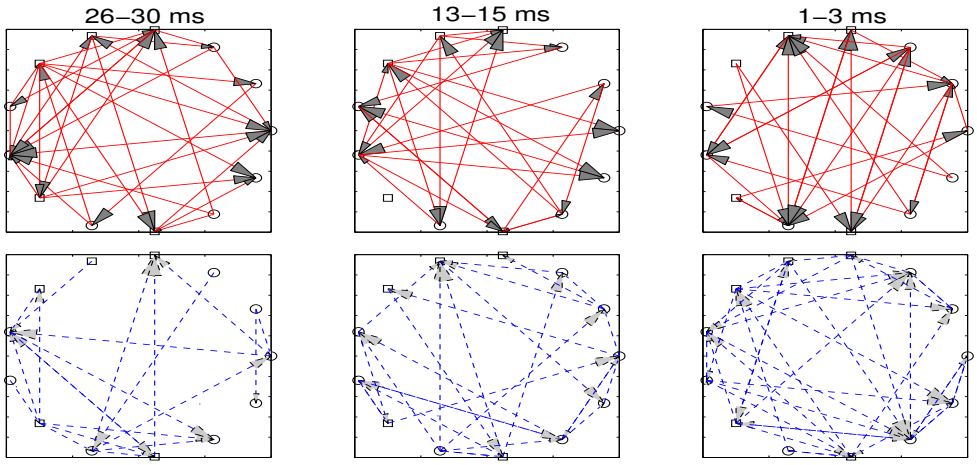

Fig. 2. Dynamic change in excitatory (top row) and inhibitory (bottom row) functional connectivity among 13 motor neurons during reaching movement (from successful trials). Circles and squares represent RS and FS cells, respectively. Uni- or bi-directional arrow indicates directional statistical dependence between cells, with solid/dashed lines representing excitatory /inhibitory connections. The title indicates the past spiking history window where the functional connectivity is inferred using all recordings.

42 trials were divided into 21 successful and 21 unsuccessful trials, respectively.

A total of 13×9 spiking history windows were empirically chosen (each cell contains 9 preceding history of spike counts), and $dim(\boldsymbol{\alpha})$ is 118 for each neuron. The empirically selected 9 firing history windows (up to preceding 40 ms) consist of the spike counts in the past 1-3, 4-6, 7-9, 10-12, 13-15, 16-20, 21-25, 26-30, 31-40 ms, respectively. Because neuronal spiking activity is *conditionally* independent (Eq. 3), individual neurons were fit with separate point process GLMs, followed by a KS test. For each neuron, the regularization coefficient ρ was selected by cross-validation. As an example, we show a few snapshots of dynamic change in functional connectivity between 13 motor neurons during reaching movement (Fig. 2). Figure 3 also shows the estimation results of one cell (#2) during baseline and reaching periods for both successful trials. Both KS plots fall within 95% confidence bounds, indicating that the point process model provides a good description of the spike trains. In this example, it also appears that during baseline, the spiking dependence coefficients between many cell pairs are close to 0 (e.g., 3→2, 7→2, 11→2, 13→2). The same cell pairs show more significant nonzero connectivity during reaching movement, suggesting that the neuronal interactions between these cell pairs are task-related. In order to characterize the time-varying functional connectivity among ensemble neurons, the excitatory (E) and inhibitory (I) connectivity ratio was computed within specific temporal windows. The ratio was defined as the total number of significant nonzero (positive or negative) coefficients against the total number of pairwise connections. The result on performance-related reaching movement is shown in Fig. 4. The average (E+I) connectivity ratio is about 0.35 over time. A close examination of Fig. 4 reveals that the number of excitatory and inhibitory connections are quite balanced (in both successful and unsuccessful trials), indicating the state of neuronal network is balanced by both excitation and inhibition. However, during successful trials, the numbers of excitatory and inhibitory connections appear more synchronized (in the same mode of increase or decrease in connectivity ratio) than in unsuccessful trials. Cell pairs were further classified according to subtypes (RS-FS, RS-RS, FS-FS) in order to investigate the incidence of functional interactions in these three groups. It was found that in both baseline and reach conditions, FS-FS pairs were most likely to display significant temporal spiking dependency, followed by RS-FS and RS-RS pairs. This observation was consistent with the findings discussed in [4]. In addition, when the total number of significant temporal spiking dependent events was compared between the baseline and reaching conditions, more interactions occurred during reaching in all three cell-pair groups, as well as for both successful and unsuccessful trials. This is in agreement with descriptions of task-related correlated neural activity [13]. In comparing skill-related successful vs. unsuccessful trials for each cell (data not shown), it was common to observe opposite excitatory vs. inhibitory effects among some cell pairs (Fig. 5). This phenomenon was seen in both RS and FS cells. This provides "statistical" evidence that the strength, timing and pseudo-postsynaptic effects of functional interactions between task-related cell pairs may play a role in coding for performance-related skill.

V. Discussion

We have developed a regularized point process GLM for characterizing functional connectivity of ensemble motor neurons during a reaching movement task. The introduction of "temporal smoothness" regularization into the model is important in that first, it *significantly* reduces the variance of the parameter estimate (due to space limit, we cannot show the non-regularized estimation results), thereby improving the generalization ability for the unseen spike train data in cross-validation (see e.g., Fig. 6); and second, it imposes a physiologically plausible constraint (as a prior) on the solution, making the interpretations of our results more meaningful.

In our preliminary analysis, we observed a dynamic temporal spiking dependency (beyond the standard 2nd-order cross-correlogram analysis) within M1 neuronal ensembles (Figs. 2-4). It was also found that significant (nonzero) statistical dependence between neuron pairs were seen more common during task performance, but not during the baseline period (Fig. 3). Interestingly, while the outcome of task performance did not appear to affect the incidence of functional interactions, the differences observed in neuronal interactions during successful and successful trials suggest that temporal coding in ensemble neurons may influence task performance.

The above findings provide a promising direction in interpreting the physiology and temporal coding among recorded motor neurons. An in-depth statistical analysis for more spike train data is currently under investigation. With further analysis, we hope that the statistical analysis within the regularized point process GLM framework

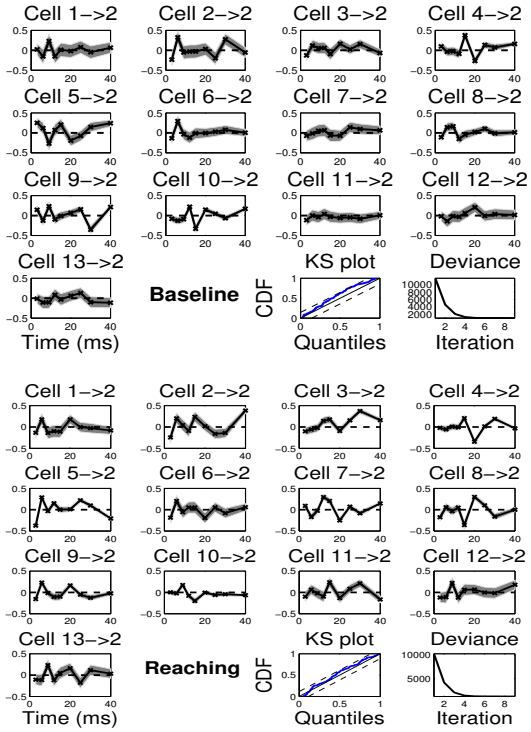

Fig. 3. An example of estimated GLM coefficients $\{\alpha_{c,k}\}$ for 9 history-dependent components from all 13 cells during baseline (top) and reaching movement (bottom). A→B assumes cell A triggers the target cell B firing with a unidirectional spiking dependence. The KS plot and the deviance convergence curve are also shown at the last two panels in each case. Shaded areas represent 95% confidence bounds.

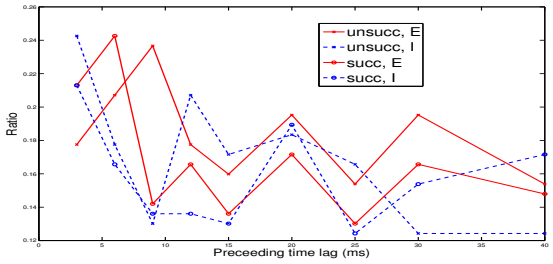

Fig. 4. The time-varying significant nonzero connectivity ratio for both excitatory (E) and inhibitory (I) connections during reaching movement within successful and unsuccessful trials.

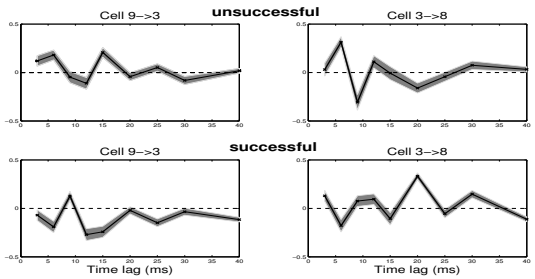

Fig. 5. Two representative examples of estimated cell pair interactions between unsuccessful and successful trials during reaching movement.

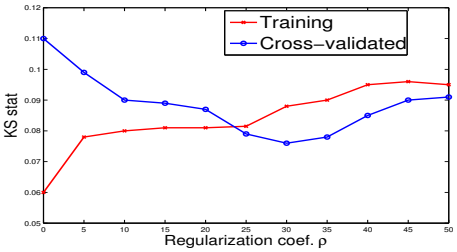

Fig. 6. The training and cross-validated KS statistics vs. the regularization coefficient ρ (cell #2, among 21 successful trials during reaching movement). In this case, the suboptimal ρ is chosen to be 30 (for $\gamma = 0.5$).

may shed some light on discovering the functional connectivity among ensemble neurons. More specifically, this may lead to developing computational tools which allow to make inferences regarding task performance based on the nature of neuronal interactions.

REFERENCES

[1] E. N. Brown, R. Barbieri, U. T. Eden, and L. M. Frank, "Likelihood methods for neural data analysis," in J. Feng (Ed.) *Computational Neuroscience: A Comprehensive Approach*, CRC Press, 2003.

[2] E. N. Brown, R. E. Kass, and P. P. Mitra, "Multiple neural spike train data analysis: state-of-the-art and future challenges," *Nature Neurosci.*, vol. 7, pp. 456–461, 2004.

[3] E. S. Chornoboy, L. P. Schramm, and A. F. Karr, "Maximum likelihood identification of neural point process systems," *Biol. Cybern.*, vol. 59, pp. 265–275, 1988.

[4] C. Constantinidis and P. S. Goldman-Rakic, "Correlated discharges among putative pyramidal neurons and interneurons in the primate prefrontal cortex," *J. Neurophy.*, vol. 88, pp. 3487–3497, 2002.

[5] G. Czanner, U. T. Eden, S. Wirth, M. Yanike, W. A. Suzuki, and E. N. Brown, "Analysis of between-trial and within-trial neural spiking dynamics," *J. Neurophy.*, vol. 99, pp. 2672–2693, 2008.

[6] M. Hebiri, "Regularization with the smooth-lasso procedure," http://arxiv.org/abs/0803.0668, 2008.

[7] P. Komarek, "Logistic regression for data mining and high-dimensional classification", TR-04-34, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, 2004.

[8] J. Kulkarni and L. Paninski, "Common-input models for multiple neural spike-train data," *Network: Computation in Neural Systems*, vol. 18, no. 4, pp. 375–407, 2007.

[9] D. Q. Nykamp, "A mathematical framework for inferring connectivity in probabilistic neuronal networks," *Math. Biosci.*, vol. 205, pp. 204–251, 2006.

[10] M. Okatan, M. A. Wilson, and E. N. Brown, "Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity," *Neural Comput.*, vol. 17, pp. 1927–1961, 2005.

[11] Y. Pawitan, *In All Likelihood: Statistical Modelling and Inference Using Likelihood*. New York: Oxford Univ. Press, 2001.

[12] J. W. Pillow, J. Shlens, L. Paninski, A. Sher, A. M. Litke, E. J. Chichilnisky, and E. P. Simoncelli, "Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature*, vol. 454, pp. 995–1000, 2008.

[13] D. F. Putrino. *Single unit and correlated neural activity observed in the cat motor cortex during a reaching movement*. unpublished PhD thesis, Univ. Western Australia, 2009.

[14] I. H. Stevenson, J. M. Rebesco, N. G. Hatsopoulos, Z. Haga, L. E. Miller, and K. P. Körding, "Bayesian inference of functional connectivity and network structure from spikes," *IEEE Trans. Neural Syst. Rehab. Eng.*, in press.

[15] R. Tibshirani, M. Saunders, S. Rosset, J. Zhu, and K. Knight, "Sparsity and smoothness via the fused lasso," *J. Roy. Soc. B*, vol. 67, pp. 91–108, 2005.

[16] W. Truccolo, U. T. Eden, M. Fellow, J. D. Donoghue, and E. N. Brown, "A point process framework for relating neural spiking activity to spiking history, neural ensemble and covariate effects," *J. Neurophy.*, vol. 93, pp. 1074–1089, 2005.