Practical Considerations of a BMI Application for Detecting Acute Pain Signals

Sile Hu, Zhengdong Xiao, Qiaosheng Zhang, Louise Urien, Jing Wang and Zhe Chen

***Abstract*— Brain-machine interfaces (BMIs) have been an important research area in closed-loop neuroscience and neuroengineering. In real-time neuroscience applications, many issues require special consideration, such as trial variability, spike sorting noise or multi-unit activity. For a BMI application of detecting acute pain signals, we discuss several practical issues in BMI applications and propose a new approach for change-point detection based on ensembles of independent detectors. Motivated from unsupervised ensemble learning, the "ensembles of change-point detectors" (ECPDs) combine the decision results from multiple independent detectors, which may be trained from data recorded at different trials or derived from different methodologies. The goal of ECPDs is to reduce the detection error (in terms of false negative and false positive rates) in online BMI applications. We validate our method using computer simulations and experimental recordings from freely behaving rats.**

***Index Terms*— Brain-machine interface, change point detection, ensemble learning, ensembles of change-point detectors, acute pain.**

I. Introduction

An important problem in closed-loop neuroscience experiments is to quickly identify abrupt changes in neural ensemble spike activity induced by external stimuli or internal changes in brain state. In real-time or closed-loop BMI applications, the challenge is to design online detection algorithms that can quickly and reliably detect the change points. Here, we consider such a BMI application for detecting acute pain signals [10]. Previously, we have designed several algorithms for testing the experimental recording off-line [5], [6]. For BMI applications, we need to consider several critical issues: (i) limited sample size for training the model; (ii) data nonstationarity or trial variability induced by behavior or adaptation; (iii) multiunit activity corrupted by spike sorting noise [9], [8], which affects the overall signal-to-noise ratio (SNR). It is worth pointing out that model estimation involves a non-global optimization procedure (i.e., multiple local maxima and sensitivity to initial conditions), and the convergence issue of model identification is more pronounced while dealing with a complex model (i.e. large number of neurons) and small data sample size. The performance of change-point detection in BMI applications is assessed by the detection speed and accuracy. We have addressed the detection speed issue in other published work [11]. Here we focus our effort on the detection accuracy (sensitivity and specificity), assessed by the false positive (FP) and false negative (FN) rates.

In this paper, motivated from the idea of ensemble learning, we propose a novel framework for change-point detection. Specifically, we construct a set of independent "weak" change-point detectors, from which we establish a meta detector based on optimal decision rules to improve the detection results. We investigate the practical issues in the context of BMI application for detecting acute pain signals based on neuronal ensemble recordings. Our method is validated using computer simulations and experimental recordings from freely behaving rats. Through computer simulations, we systematically investigate the impact of cell specificity (i.e., positively or negatively modulated neuronal response to pain stimuli), SNR, and trial variability on the performance of *ensembles of change-point detectors* (ECPDs). Finally, we conclude the paper with discussions on future work.

S. Hu and Z. Xiao are with the Department of Instrument Science and Technology, Zhejiang University, Hangzhou, Zhejiang, China, and are also with the Department of Psychiatry, New York University School of Medicine, New York, NY, USA.

Q. Zhang, L. Urien and J. Wang are with the Department of Anesthesiology, Perioperative Care, and Pain Medicine, New York University School of Medicine, New York, NY, USA.

Z. Chen is with the Department of Psychiatry, and Department of Neuroscience and Physiology, New York University School of Medicine, New York, NY, USA. (Correspondence email: zhe.chen3@nyumc.org)

II. Methods

A. Model-based Method

Our model-based method for change-point detection is based on state space analysis [3], [5]. Let $\mathbf{y}_k = [y_{1,k}, \ldots, y_{C,k}]^\top$ denote a C-dimensional population vector, with each element consisting of the neuronal spike count at the k-th time bin (bin size Δ). Let the latent univariate variable $z_k \in \mathbb{R}$ represents an unobserved common input that drives neuronal ensemble spiking activity.

1) Poisson Linear Dynamical System: We consider a Poisson linear dynamical system (PLDS) [5], [16], where the spike activity of a population of C neurons are assumed to be drawn from the following generative model:

$$z_k = az_{k-1} + \epsilon_k \quad (1)$$

$$\mathbf{y}_k \sim \texttt{Poisson}\Big(\exp(\mathbf{c}z_k + \mathbf{d})\Delta\Big) \quad (2)$$

where the state equation (1) describes a first-order autoregressive (AR) model ($0 < |a| < 1$) driven by a zero-mean Gaussian noise process $\epsilon_k \in \mathcal{N}(0, \sigma_\epsilon^2)$. The parameters $\mathbf{c}$ and $\mathbf{d}$ are unconstrained. We use an expectation-maximization (EM) algorithm [16], [5] to estimate the unknown state variables $z_{1:T}$ and parameters $\mathbf{\Theta} = \{a, \mathbf{c}, \mathbf{d}, \sigma_\epsilon\}$ from a set of observations $\mathbf{y}_{1:T}$.

2) Online Recursive Filtering: In online BMI applications, once the model parameters are identified, we use a recursive (forward) filter to estimate the latent state variable [5], [6]:

$$\hat{z}_{k|k-1} = a\hat{z}_{k-1|k-1} \quad (3)$$

$$Q_{k|k-1} = a^2 Q_{k-1|k-1} + \sigma_\epsilon^2 \quad (4)$$

$$\hat{\mathbf{y}}_{k|k-1} = \exp(\mathbf{c}\hat{z}_{k|k-1} + \mathbf{d})\Delta \quad (5)$$

$$Q_{k|k}^{-1} = Q_{k|k-1}^{-1} + \mathbf{c}^\top diag(\hat{\mathbf{y}}_{k|k-1})\mathbf{c} \quad (6)$$

$$\hat{z}_{k|k} = \hat{z}_{k|k-1} + Q_{k|k}\mathbf{c}^\top(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \quad (7)$$

where $Q_{k|k} = \mathrm{Var}[\hat{z}_{k|k}]$ denotes the filtered state variance. In our online filtering algorithm, we set the initial condition $\hat{z}_{0|0} = 0$ and $Q_{0|0}$ from the previously trained model.

3) Change-Point Detection: From the online filtered estimate $\hat{z}_{k|k}$, we compute the Z-score related to the baseline: Z-score $= \frac{\hat{z}_{k|k} - \text{mean of } z_{\text{baseline}}}{\text{SD of } z_{\text{baseline}}}$ and convert it to probability [4], [5]:

$$P(\text{Z-score} > \xi_0) = 1 - \int_{-\infty}^{\xi_0} \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}} du \quad (8)$$

The criterion of Z-score change is determined by a critical threshold ξ_0. Using the 95% significance level ($\xi_0 = 1.65$), it is concluded that when Z-score $-$ CI > 1.65 or Z-score $+$ CI < -1.65, where the CI denotes the confidence interval: $2\sqrt{Q_{k|k}}/\text{SD of } z_{\text{baseline}}$.

B. Model-free Method

Under the Poisson assumption, we compute the baseline firing rate of a neuron, say λ_0. If we consider the threshold of significant change in firing rate (in the positive direction) as $\lambda_{th} = \lambda_0 + 3\sqrt{\lambda_0}$. We can compute the log-likelihood ratio (LLR) as

$$\text{LLR} = y \log \frac{\lambda_{th}}{\lambda_0} - (\lambda_{th} - \lambda_0) \tag{9}$$

Our model-free method uses a so-called CUSUM (cumulative sum) approach to detect a change point based on LLR [6]. Setting $S_0 = 0$, we update the cumulative sum as follows

$$S_k = \max_{c=1,\ldots,C} \Big\{ S_{c,k} \Big\} = \max_{c} \Big\{ \max \Big\{ 0, S_{c,k-1} + s_{c,k} \Big\} \Big\} \tag{10}$$

where $s_{c,k}$ is computed from the LLR (9) using $y_{c,k}$ from the c-th neuron. If the cumulative sum of the statistic S_k is above a predetermined threshold θ_0 and the trend continues more than 150 ms and the trend is monotonic, then a decision is made regarding the presence of a change point. The threshold θ_0 in the CUSUM algorithm controls the false alarm rate. An empirical choice is to use the test statistic (twofold log-likelihood) being a chi-square distribution with 1 degree of freedom: $\chi^2_{1,(1-\alpha)}$. If we set $\alpha = 0.01$, we have $\theta_0 = 0.5 \times 6.64 = 3.38$.

C. Ensembles of Change-point Detectors

Ensemble learning, in supervised or unsupervised form, has been an active research topic in machine learning [13], [7], [14], [18]. Ensemble learning is aimed to combine a set of possibly "weak" learners (predictors, classifiers, detectors, etc.) to form a more accurate "meta learner" for decision making. In unsupervised ensemble learning, the decision is made without labelled data [17], [12]. There are several differences in our current application from other unsupervised ensemble learning: (i) In ensembles of unsupervised classifiers, samples are independent, identically distributed (i.i.d.); whereas samples in our application are temporal correlated multivariate time series. (ii) In traditional ensemble learning, the predictors are derived from independent models of different families; whereas our independent detectors are derived from the same model family. (iii) We combine model-based and model-free methods or integrate decisions from two brain regions for change-point detection.

Given N independent detectors, at any time point, from (8) we can derive a probability of significance change from each detector. In general, assume that the j-th detector produces a probability of change at time k: $P(\omega_{j,k}|\mathbf{y}_{0:k})$ with a prior probability P_j, then the decision rule for $\omega_k \in \{0, 1\}$ (1 denotes a change and 0 no change) is formulated mathematically as follows:

$$\Pr(\omega_k = 1) = \prod_{j=1}^{N} P(\omega_{j,k}|\mathbf{y}_{0:k}) P_j \tag{11}$$

Similar to pattern classification [13], we can also design different rules. Here, we consider two decision rules: *greedy rule* and *majority vote rule*. The greedy rule claims a change as long as one detector predicts the change. The majority vote rule states that the class $\omega_k = 0/1$ that receives the largest number of votes is selected as the consensus or majority vote, assuming equal P_j.

D. Combining Model-based and Model-free Detection

Our model-free approach for change-point detection is a greedy method. Our previous empirical studies have shown that it tends to produce similar or faster detection than the model-based approach. However, the model-free approach is also sensitive to noise, thereby prone to the FP. Normally, to minimize the FN rate, we impose a a *necessary* condition of TP decision from the model-free method.

E. Combining Two Brain Regions

Among many brain regions, the primary somatosensory cortex (S1) and anterior cingulate cortex (ACC) are two of the most studied areas for pain-related perception [15], [20], [2]. The S1 encodes the sensory component of pain, whereas the ACC encodes the aversive component of pain. However, neither of these areas are pain-specific. Namely, they may encode other sensory or emotional responses to non-noxious stimuli (e.g., touch, auditory tone, etc). In our previous work, we found that S1 and ACC populations have different sensitivity and specificity for pain encoding [5]. In general, S1 tends to have high sensitivity for noxious stimuli, but it can also cause false alarms for non-noxious stimuli. On the other hand, ACC tends to have mixed or conjunctive coding for a variety of stimuli related to emotion, nociception and anticipation. Therefore, it is important to design specific decision rules for detecting acute pain based on their response properties.

Given simultaneous S1 and ACC ensemble recordings, we cannot use the greedy or majority vote rule. We envision two options. One option is to use different priors in (11) for S1 and ACC, but the prior probabilities may depend on specific dataset. In another option, we consider a decision rule based on a dynamic cross-correlation function (CCF) between the Z-scores Z_k of two regions:

$$CCF_k = (1-\rho) CCF_{k-1} + \rho \Big(Z_k^{\text{ACC}} \Big)^m \Big(Z_k^{\text{S1}} \Big)^n \tag{12}$$

where $0 < \rho < 1$ is a forgetting factor, $0.5 \le m, n \le 1$ are the scaling exponents (in our analysis, we set $m = n = 0.5$). The idea behind that is when two Z-scores follow a consistent trend, the CCF will increase in absolute value, otherwise it will stay around the baseline level. Similarly, we can compute Z-score of CCF relative to the baseline. When the CCF area above the twofold SD exceeds a threshold (e.g., 3), we will declare the change point.

III. Results

A. Computer Simulations

a) Trial variability: We generate population spike trains of $C = 24$ units using a PLDS model with assumed ground truth. Each trial lasts 10 s and consists of 4 pain-modulated units (all positively modulated), with randomly generated vector $\mathbf{c}$ and $\mathbf{d}$ in Eq. (2). The values $\mathbf{c}$ and $\mathbf{d}$ are set within a predefined range to keep the Poisson spike count at each time bin less than 8. The duration of pain stimulus is 2 s. To simulate trial variability, we create different configurations where the specified number of coefficients in vector $\mathbf{c}$ associated with the pain-modulated neurons are varying across trials. Without loss of generality, we define

Class 1: pain-modualted unit index $\in \{1, 2, 3, 4\}$

Class 2: non-pain-modulated unit index $\in \{5, 6\}$

Class 3: non-responsive unit index $\in \{7, 8, \ldots, \}$

For the TP experiment, we train $n = 100$ models using all 100 randomly generated trials. After that, we apply all models to detect changes in all trials, yielding an $n \times n$ matrix for each dataset. We compute the average detection results based on single detector or ECPDs using the majority vote. In this case, the testing set and the training set are the same. The detailed results are summarized in Table I and Fig. 1A. As seen, in nearly all configurations, the ECPDs based on the majority vote outperforms the single model-based detection accuracy in TP. The relative improvement degree is greater in the presence of higher trial-by-trial variability.

TABLE I

SUMMARY OF SIMULATION EXPERIMENTS (SETUP 1, $C = 24$). THE SMALLER THE NUMBER OF INVARIANT CLASS-1 UNITS, THE HIGHER TRIAL-VARIABILITY WITHIN A SPECIFIC DATA SET. DATASETS 1-4: LOW TRIAL VARIABILITY; DATASETS 5-10: MEDIUM TRIAL VARIABILITY; DATASETS 11-14: HIGH TRIAL VARIABILITY.

Dataset	# trials	#invariant Class-1 units	invariant Class-1 units	TP (single)	TP (ensemble)
1	100	3	1, 2, 3	95%	98%
2	100	3	1, 2, 4	93%	94%
3	100	3	1, 3, 4	96%	96%
4	100	3	2, 3, 4	95%	95%
5	100	2	1, 2	75%	86%
6	100	2	1, 3	79%	87%
7	100	2	1, 4	85%	88%
8	100	2	2, 3	87%	93%
9	100	2	2,4	77%	81%
10	100	2	3,4	84%	86%
11	100	1	1	52%	51%
12	100	1	2	56%	69%
13	100	1	3	67%	84%
14	100	1	4	75%	84%
15	100	0		46%	34%

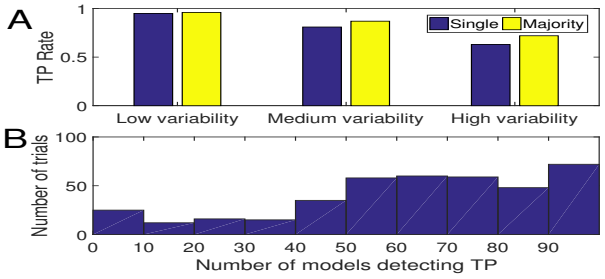

Fig. 1. (*A*) Average performance comparison between single-model detector and ECPDs using the majority vote under low (with 3 invariant units), medium (with 2 invariant units) and high (with 1 invariant unit) trial-to-trial variability. (*B*) Histogram of the number of high-variability trials (Datasets 11-14) that can be detected as TPs. Among 400 trials, each trial is tested with 100 trained models.

Among all 400 high-variability trials (Datasets 11-14), about 75% of trials benefit from the majority vote rule; namely, the trials that benefit from the majority vote are located within the right half of distribution (i.e., support > 50, Fig. 1B).

To assess the FP rate in change point detection, we use a similar setup but with fewer units ($C = 12$). In the training trials, it is assumed that two Class-2 units fire in response to a presumed non-pain stimulus (0.5 s duration) during the baseline with probability q, and these 2 Class-2 units do not overlap with the 4 Class-1 units mentioned above. We vary the probability $q \in \{0.1, 0.25, 0.5, 1\}$ and simulate $n = 100$ training and testing trials. In the testing trials, we generate population spike responses with only "positive" responses from these 2 Class-2 units, whereas the firing rates of remaining 10 units are the same as the baseline. We train 100 models from all training trials and test them on all 100 testing trials, from which the FP rate is computed using either single detector or ECPDs. As seen in Table II, ECPDs significantly reduce the FP rate based on the majority vote for a wide range of q values.

b) Noise: To investigate the impact of noise to change-point detection, we add independent Poisson noise to population spike activity across time and units. We vary the level of SNR: $10 \log_{10} \frac{\lambda_{signal}}{\lambda_{noise}}$ (note that Poisson mean is equal to variance),

TABLE II

SUMMARY OF SIMULATION EXPERIMENTS ($C = 12$). LARGER q IMPLIES A LOWER SNR IN BASELINE AMONG TRAINING TRIALS.

Dataset	prob. q	accuracy (single)	FP (single)	FP (ensemble)
16	1.0	85%	44%	25%
17	0.5	85%	27%	0
18	0.25	84%	16%	0
19	0.1	84%	9%	0

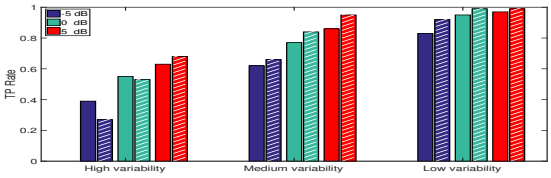

Fig. 2. Average TP rate comparison between single model detector and ECPDs (shaded) under different levels of SNR and trial variability.

where the SNR is constant across trials. We repeat the experiment in simulation setup 1 and compare the TP rate between the single-model detector and ECPDs. The results are shown in Fig. 2. As seen, ECPDs outperform the single-model detector at medium and high variabilities with various SNRs; however, the superiority of ensemble learning is lost in the case of high trial variability and very low SNR (i.e., -5 dB and 0 dB).

B. Experimental Recordings

All procedures in this study were performed in accordance with the New York University School of Medicine Institutional Animal Care and Use Committee and the NIH Guideline. Male Sprague-Dawley rats were used in all experiments. The pain stimulus was delivered by a blue (473 nm diode-pumped solid-state) laser with varying laser intensities (50-250 mW) [5]. The inter-trial intervals between consecutive laser stimulations were at least 1 min. Animals freely explored in a plastic chamber of size $38 \times 20 \times 25$ cm^3. One video camera (120 frame per second) was used to continuously record the animal's behavior. We used custom tetrode/stereotrode arrays (a total of 32 channels) or two bundled silicon probes (64 channels) to record neural activity from the rat ACC or S1 areas, or simultaneously). Using a Plexon (Dallas, TX) data acquisition system, spikes were thresholded from high-passed (>300 Hz) local field potentials (LFPs). The detected spikes were further sorted using online spike-sorting software (Plexon). For the illustration purpose, we validate our method using a few experimental recording sessions.

The first recording session consists of 32 ACC units and 25 experimental trials of 150 mW laser stimulation. To assess trial-by-trial variability, we estimate one model from one trial and then apply 25 models to test all 25 trials, yielding a binary 25-by-25 matrix on detection results, with 1 and 0 representing TP and FN, respectively (Fig. 3A). Furthermore, to emulate the online BMI application, we use only 5 models trained from preceding 5 trials and test the current trial. An improvement in TP (averaged across $25 - 5 = 20$ trials) is shown from single model detector to ECPDs using the majority vote (Fig. 3B). Overall, we see a slight average improvement. An example is shown in Fig. 4, in which 4 out of 5 models detect the actual change induced by the pain stimulus..

The second recording session consists of 26 ACC units with a 250 mW laser stimulation (during a conditioning experiment) [19]. Without going into experimental details, we use one experimental

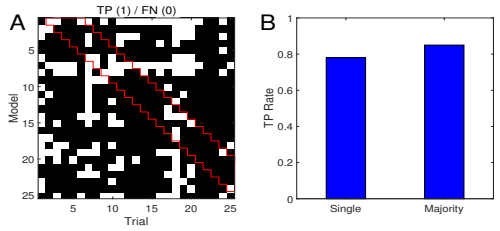

Fig. 3. (*A*) Detection results of the true positive (TP) matrix. Each entry of the matrix denote the detection result from one model (row) applied to one trial (column). Black/white color denotes TP/FN. Banded diagonal in red denotes the results using models trained from preceding 5 trials. (*B*) Comparison of averaged TP rate (averaged across 20 trials).

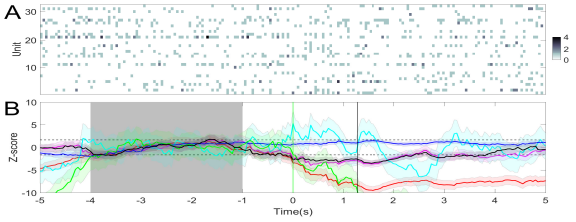

Fig. 4. (*A*) A trial example of spike count observations of $C = 32$ ACC units. Time 0 denotes the onset of 150 mW laser stimulation. (*B*) Z-scores derived from 6 PLDS models trained from the current (n, *red*) and preceding trials ($n-1$, *magenta*; $n-2$, *cyan*; $n-3$, *blue*; $n-4$, *green*; $n-5$, *black*). Vertical red and black lines denote the laser onset and paw withdrawal, respectively. Horizontal dashed lines denote the ± 1.65 threshold.

trial to illustrate the benefit of combining model-based and model-free approaches in change-point detection when there are very few Class-1 units. As shown in Fig. 5, the model-based method fails to detect the change due to low number of Class-1 units (or high certainty in the Z-score estimate). However, the model-free method is able to detect the change. It is found that combining these two methods helps improve the TP rate in this session.

Finally, the third and fourth recording sessions consist of simultaneously recorded ACC and S1 units of rats subject to double laser simulations. The 50 mW laser stimulation is a negative control for pain. As illustrated in one trial example of 50 mW stimulation (Fig. 6), the ACC region detects two change points; in contrast, the CCF from two ensembles avoids the FP. The result is also robust to the choice of ρ. In another trial example of 250 mW laser stimulation (Fig. 7), S1 detects the change, whereas ACC does not. In contrast, the CCF detects the TP for a wide range of forgetting factors. Overall, combining ACC and S1 can improve the detection accuracy in many experimental trials.

IV. Discussion and Conclusion

In our proposed ECPDs, the individual change-point detectors are estimated from different trials. Our previous BMI system consists of two parallel CPU threads, one used for model estimation, and another used for recursive filtering [10]. However, when the numbers C and N are very large, the increasing computational cost may induce undesirable time delay. To allocate more CPU computing resources, we may consider using an optimized digital signal processing (DSP) board dedicated to model estimation.

In addition to ensemble spike activity, LFPs may also reveal important information about pain signals, in both time and frequency domains [22]. However, LFP signals are noisy due to potential corruption of movement artifacts in freely behaving animals. Once artifact rejection is achieved, we can first design an independent

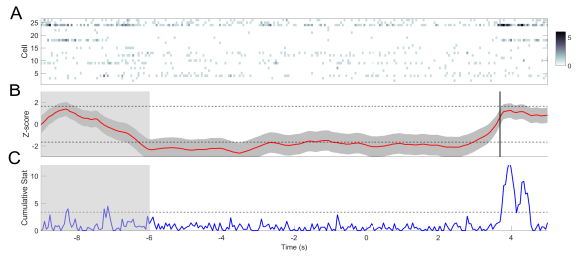

Fig. 5. (*A*) A trial example of spike count observations of $C = 26$ ACC units. (*B*) Z-score curves from one PLDS model. Horizontal dashed lines denote the ± 1.65 threshold. (*C*) Cumulative statistic in the model-free method. Horizontal dashed line shows the threshold 3.38.

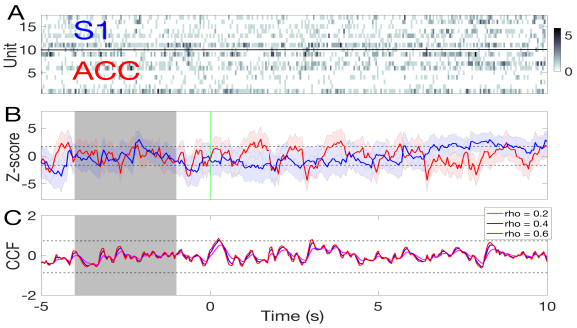

Fig. 6. (*A*) A trial example of rat ACC ($C = 10$) and S1 ($C = 7$) population spike count observations with a 50 mW laser stimulation at time 0. (*B*) Z-score curves derived from ACC (*red*) and S1 (*blue*). Horizontal dashed lines denote the ± 1.65 threshold. Vertical red denotes the laser onset. (*C*) Z-scored cross-correlation function (CCF) from both areas based on different values of forgetting factor ρ; $m = n = 0.5$.

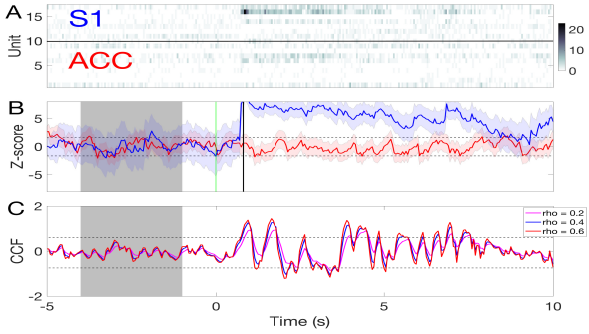

Fig. 7. (*A-C*) Same as Fig. 6, but for 250 mW laser stimulation at time 0.

(supervised or unsupervised) change-point detector for pain signals; then we can adapt the strategy of ECPDs and integrate the information from spikes and LFP for detecting pain signals.

In summary, we propose an unsupervised ensemble learning framework for change-point detection. The general framework may accommodate multiple trials, multiple brain regions and multiple detection methods. In the future work, we will further investigate optimal decision rules within this ensemble detection framework and test on more experimental recordings.

Acknowledgments

The work was supported by the US NSF-CRCNS grant IIS-130764 (Z.C.), NIH grants R01-NS100016 (Z.C., J.W.), and R01-GM115384 (J.W.).

REFERENCES

[1] L. Becerra, E. Navratilova, F. Porreca and D. Borsook, "Analogous responses in the nucleus accumbens and cingulate cortex to pain onset (aversion) and offset (relief) in rats and humans," *J. Neurophysiol.*, vol. 110, no. 5, pp. 1221–1226, 2013.

[2] M. C. Bushnell, M. Ceko and L. A. Low, "Cognitive and emotional control of pain and its disruption in chronic pain," *Nat. Rev. Neurosci.*, vol. 14, pp. 502–511, 2013.

[3] Z. Chen, editor, *Advanced State Space Methods for Neural and Clinical Data*, Cambridge Univ. Press, 2015.

[4] Z. Chen and J. Wang, "Statistical analysis of neuronal population codes for encoding acute pain," *Proc. IEEE ICASSP*, pp. 829–833, 2016.

[5] Z. Chen, Q. Zhang, A. P. S. Tong, T. R. Mander and J. Wang, "Deciphering neuronal population codes for acute thermal pain," *J. Neural Engr.*, vol. 14, no. 3, p. 036023, 2017.

[6] Z. Chen, S. Hu, Q. Zhang and J. Wang, "Quickest detection for abrupt changes in neuronal ensemble spiking activity using model-based and model-free approaches," *Proc. IEEE Neural Engineering Conf. (NER'17)*, pp. 481–485, 2017.

[7] T. G. Dietterich, "Ensemble methods in machine learning", in N. El Gayar, J. Kittler, and F. Roli, eds. *Multiple Classifier Systems (9th International Workshop, MCS'2000)*, Springer, 2000.

[8] G. W. Fraser, S. M. Chase, A. Whitford, and A. B. Schwartz, "Control of a brain-computer interface without spike sorting," *J. Neural Engr.*, vol. 6, pp. 055004, 2009.

[9] I. N. Goodman and D. H. Johnson, "Information theoretic bounds on neural prosthesis effectiveness: The importance of spike sorting," *Proc. IEEE ICASSP08*, pp. 5204-5207, 2008.

[10] S. Hu, Q. Zhang, J. Wang and Z. Chen, "A real-time rodent neural interface for deciphering acute pain signals from neuronal ensemble spike activity," *Proc. 51st Asilomar Conf. on Signals, Systems, and Computers*, 2017.

[11] S. Hu, Q. Zhang, J. Wang and Z. Chen, "Real-time particle filtering and smoothing algorithms for detecting abrupt changes in neural ensemble spike activity," *J. Neurophysiol.*, vol. 119, 2018. Online https://doi.org/10.1152/jn.00684.2017.

[12] A. Jaffe, B. Nadler, and Y. Kluger, "Estimating the accuracies of multiple classifiers without labeled data," *Proc. AISTAT'15*, 2015.

[13] J. Kittler, M. Hatef, R. P. W. Duin and J. Matas, "On combining classifiers," *IEEE Trans. Patt. Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, 1998.

[14] L. I. Kuncheva, *Combining Pattern Classifiers: Methods and Algorithms*, Wiley, 2004.

[15] C. C. Kuo and C. T. Yen, "Comparison of anterior cingulate and primary somatosensory neuronal responses to noxious laser-heat stimuli in conscious, behaving rats," *J. Neurophysiol.* vol. 94, pp. 1825–1836, 2005.

[16] J. H. Macke, L. Buesing and M. Sahani, "Estimating state and parameters in state space models of spike trains," in Z. Chen (ed.), *Advanced State Space Methods for Neural and Clinical Data* (Chap. 6, pp. 137–159). Cambridge Univ. Press, 2015.

[17] F. Parisi, F. Strino, B. Nadler, and Y. Kluger, "Ranking and combining multiple predictors without labeled data," *Proc. Natl. Acad. Sci. USA*, vol. 111, no. 4, pp. 1253–1258, 2014.

[18] V. C. Raykar, Y. Shipeng, L. H. Zhao, G. H. Valdez, C. Florin, L. Bogoni and L. Moy, "Learning with crowds," *J. Machine Learning Res.*, vol. 11, pp. 1297–1322, 2010.

[19] L. Urien, Z. Xiao, J. Dale, E. Bauer, Z. Chen and J. Wang, "Rate and temporal coding mechanisms in the anterior cingulate cortex for pain anticipation," *Sci. Rep.*, to appear.

[20] B. A. Vogt, "Pain and emotion interactions in subregions of the cingulate gyrus," *Nat. Rev. Neurosci.*, vol. 6, pp. 533–544, 2005.

[21] Q. Zhang, T. R. Mander, A. P. Tong, R. Yang, A. Garg, E. Martinez, H. Zhou, J. Dale, A. Goyal, L. Urien, G. Yang, Z. Chen, and J. Wang, "Chronic pain induces generalized enhancement of aversion," *eLife*, vol. 6, pp. e25302, 2017.

[22] Q. Zhang, Z. Xiao, C. Huang, S. Hu, P. Kulkarni, E. Martinez, A. P. S. Tong, A. Garg, H. Zhou, Z. Chen, and J. Wang, "Local field potential decoding of the onset and intensity of acute thermal pain in rats," *Sci. Rep.*, under review.