Published in final edited form as:

J Neural Eng. 2017 June ; 14(3): 036023. doi:10.1088/1741-2552/aa644d.

Deciphering neuronal population codes for acute thermal pain

Zhe Chen^{1,2}, Qiaosheng Zhang³, Ai Phuong Sieu Tong^{4,5}, Toby R Manders^{3,6}, and Jing Wang^{2,3}

¹Department of Psychiatry, New York University School of Medicine, New York, NY 10016, United States of America

²Department of Neuroscience and Physiology, New York University School of Medicine, New York, NY 10016, United States of America

³Department of Anesthesiology, Perioperative Care and Pain Medicine, New York University School of Medicine, New York, NY 10016, United States of America

⁴Department of Biomedical Engineering, Columbia University, New York, NY 10019, United States of America

⁵University of Washington School of Medicine, Seattle, WA 98195, United States of America

⁶Washington University School of Medicine in St. Louis, St. Louis, MO 63110, United States of America

Abstract

Objective—Pain is defined as an unpleasant sensory and emotional experience associated with actual or potential tissue damage, or described in terms of such damage. Current pain research mostly focuses on molecular and synaptic changes at the spinal and peripheral levels. However, a complete understanding of pain mechanisms requires the physiological study of the neocortex. Our goal is to apply a neural decoding approach to read out the onset of acute thermal pain signals, which can be used for brain-machine interface.

Approach—We used micro wire arrays to record ensemble neuronal activities from the primary somatosensory cortex (S1) and anterior cingulate cortex (ACC) in freely behaving rats. We further investigated neural codes for acute thermal pain at both single-cell and population levels. To detect the onset of acute thermal pain signals, we developed a novel latent state-space framework to decipher the sorted or unsorted S1 and ACC ensemble spike activities, which reveal information about the onset of pain signals.

Main results—The state space analysis allows us to uncover a latent state process that drives the observed ensemble spike activity, and to further detect the ‘neuronal threshold’ for acute thermal pain on a single-trial basis. Our method achieved good detection performance in sensitivity and specificity. In addition, our results suggested that an optimal strategy for detecting the onset of

Author contributions. ZC and JW conceived and designed the experiments; QZ, APST, TRM performed the experiments; ZC analyzed the data, with assistance from QZ and APST; ZC and JW wrote the paper, with contributing comments from the other authors.

acute thermal pain signals may be based on combined evidence from S1 and ACC population codes.

Significance—Our study is the first to detect the onset of acute pain signals based on neuronal ensemble spike activity. It is important from a mechanistic viewpoint as it relates to the significance of S1 and ACC activities in the regulation of the acute pain onset.

Keywords

acute pain; population codes; primary somatosensory cortex; anterior cingulate cortex; neural decoding

1. Introduction

Pain is the leading cause of disability and a major contributor to health care costs. In basic and translational pain research, animal models are pivotal for understanding the mechanisms of pain (Mogil 2009, Jaggi *et al* 2011). Research over the last 50 years has resulted in a better understanding of spinal and peripheral mechanisms for pain. Nevertheless, brain circuits that regulate pain, especially the affective component of pain, remain poorly understood. Such understanding, however, is vital to the knowledge of how the brain processes sensory and affective information, which could help lead to novel therapeutic strategies.

Clinically, pain is primarily assessed by various forms of self-report. However, self-reports are unavailable or unreliable in certain patient populations, such as pediatric and geriatric patients as well as patients with cognitive impairment. Furthermore, self-reports tend to be subjective and can be affected by a wide range of concurrent emotional and cognitive processes. Thus, a relatively unbiased read-out from neural activities in the cortex may provide supplementary and in some cases the only measurements of pain. There is ongoing work in this direction from the human imaging field (Becerra and Borsook 2008, Baliki *et al* 2012, Wagner *et al* 2013, Ung *et al* 2014). However, imaging lacks the capability for providing data of high temporal resolution, neither provide direct information of neural activities. In contrast, electrophysiological recordings provide much better temporal resolution for acute pain signals.

Neuroimaging studies in humans have suggested that among many brain regions, the primary somatosensory cortex (S1) and anterior cingulate cortex (ACC) are two of the most studied cortical areas for pain-related perception. Notably, S1 and ACC represent two diverse stages of input elaboration and have different roles and timings in processing pain signals. The S1 has been thought to indicate the *sensory-discriminative* component of pain (Vierck *et al* 2013), whereas the ACC has been widely hypothesized to encode the *affective-motivational* component of pain (Sewards and Swards 2002, Vogt 2005, Bushnell *et al* 2013, Fuchs *et al* 2014). Specifically, S1 receives various projections from spinal and brain stem neurons that receive their inputs from myelinated afferents. Myelinated nociceptive input to S1 involves transmission of activity of L1–L4 neurons (from four lumbar nerve segments) in the spinal dorsal horn via the spinothalamic tract and subsequently via the contralateral thalamus. In contrast, ACC receives primarily its afferent axons from the

medial thalamus, including the intralaminar and midline thalamic nuclei, with additional input from other cortical areas including the prefrontal cortex (PFC). Therefore, ACC has a multi-functional role in modulation of affective pain responses, whereas S1 provides more specific information regarding the sensory component of acute pain. Previously, several electrophysiological studies have established that activities in the S1 and ACC are altered by pain stimuli (Kuo and Yen 2005, Zhang *et al* 2011, Li *et al* 2014). However, very little effort has been devoted to decoding analysis or onset detection of pain signals. A great computational challenge is to identify relevant neural ensembles that drive pain behavior. The identification of such neuronal ensembles could effectively provide a ‘neural code’ for pain. It is even more important to design algorithms that enable us to read out the neural code for pain based on the ensemble spike activity. Such a neural signature for pain can have profound clinical implications in cases when patients are unable to provide verbal measures of pain.

Acute pain is defined with pain with a sharp and well-defined onset with relatively short duration. We use a rat model to study acute thermal pain, where a laser with controlled intensity is delivered to the animal to provide a nociceptive stimulus. Paw withdrawal is used to indicate the nociceptive response. We record *in vivo* ensemble spike activities from the S1 and ACC in freely behaving rats. We develop a latent-variable modeling framework for detecting the onset of acute pain based on S1 or ACC ensemble spike activities in response to noxious laser stimuli.

Stimulus detection can be viewed as a special case of decoding analysis. In the neuroscience field, the traditional decoding paradigm is built upon establishing an encoding model first based on behavioral measures or supervised labels, such as the generalized linear model (GLM) (Brown *et al* 1998) or support vector machine (Ambard and Rotter 2012). Our work distinguishes from the previous literature. Unlike previous studies that correlated neural firing with pain intensities (Zhang *et al* 2011), our study takes on a unique task of detecting the onset of pain using an unsupervised learning approach. Specifically, our approach employs state space analysis (Chen 2015) and variational inference for a state space model (SSM). The latent state represents a common input that drives Poisson-distributed population spike activity in time. Therefore, detecting the relative change in the latent state reveals information of the onset of acute pain signals. The state space approach allows us to identify the smoothed latent variable, the noise statistics, and the relative strength of each neuron modulated by the latent variable. We verify our method with experimentally recorded ensemble spike data and assess its detection sensitivity and specificity.

2. Materials and methods

2.1. Electrophysiology and data recording

2.1.1. Animals—All procedures in this study were performed in accordance with the New York University School of Medicine (NYUSOM) Institutional Animal Care and Use Committee and the National Institutes of Health (NIH) Guide for the Care and Use of Laboratory Animals (publication number 85–23) to ensure minimal animal use and discomfort. Male Sprague-Dawley rats were purchased from Taconic Farms and kept at Mispro Biotech Services Facility in the NYC Alexandria Center for Life Science, at room

temperature, with controlled humidity, 12 h (6:30 a.m.–6:30 p.m.) light/ dark cycle, and ad libitum access to food and water. Animals (300–350 g at implantation and testing) were given on average 7 d to adjust to the new environment before surgery. A 7 d post-surgical recovery period was given before the onset of any experiments.

2.1.2. Electrode implantation and surgery—Tetrodes were constructed from four twisted 12.7 μm polyimide coated microwires (Sandvik) and mounted in an 8 tetrode VersaDrive (Neuralynx). Electrode tips were plated with gold to reduce electrode impedances to 100–500 k Ω at 1 kHz. Rats were anesthetized with isoflurane (1.5–2%). The skull was exposed and a 2.5 mm-diameter hole was drilled above the target region. A durotomy was performed before tetrodes were slowly lowered unilaterally into the S1 or ACC with the stereotaxic apparatus. Coordinates for ACC implants were: anteroposterior (AP) 2.7 mm, mediolateral (ML) 0.6 mm, and dorsoventral (DV) 1.4 mm, with tetrode tips angled 10° toward the midline. Coordinates for hind limb S1 implants were: AP 1.5 mm, ML 3 mm, and DV 0.5 mm (figure 1). The drive was secured to the skull with screws with dental cement.

2.1.3. Laser stimulation—Noxious stimulation via a 473 nm blue diode-pumped solid-state laser (Shanghai Dreams Lasers Technology Co., LTD.) was applied 1 mm from the plantar surface of the hind paw *contralateral* to the brain recording site in freely moving rats. The laser was turned on by a transistor-to-transistor (TTL) pulse generator (Doric) until paw withdrawal was observed. The fiber output power was calibrated by compact power and energy meter console (PM100D, Thorlabs) at the beginning of every recording day. The diameter of the fiber tip was 1 mm. Neuronal activity and the onset of noxious laser stimulation were simultaneously recorded with an acquisition board (Open Ephys). A 30 frame-per-second video camera (Sony) was used to record the onset of pain, indicated by rapid withdrawal. The average stimulus duration at a laser power of 250 mW was <2 s. The duration of the stimulus was related to paw withdrawal, as the act of withdrawal removed the paw from the heat source. The latency to withdrawal depends on a number of factors including (i) power of laser, (ii) heat conduction through the tissue, (iii) neural transmission through the peripheral nociceptive neuron and the speed of spinal reflex, and (iv) signal transmission through the neuro-muscular junction. There is variability in each of these factors, and therefore there is also variability in withdrawal latency and stimulus duration.

All recording sessions consisted of multiple trials with variable inter-trial intervals (most >45 s, except for Dataset rat2-S1-2 that used >10 s) using one type of stimulation, with at least 1 h in between sessions. To avoid sensitization due to repeated stimulation and testing, we have carefully monitored both neural and behavioral responses to pain stimuli over the period of all the trials within each experiment, and have taken extra care to minimize any potential algnesia induced by overstimulation. We did not identify any changes in behavior after repeated stimulation. Neither did we observe any physical injury or inflammatory reaction in the paw during our experiment. In each session, we did not see any systematic trend in withdrawal latency between the first 5 and last 5 trials (two examples in figure 2), nor did we observe a faster onset of spike responses at single neuron levels over the period

of 30–40 trials. Therefore, we do not believe that sensitization was a factor in our experiments.

2.1.4. Temperature measurement—We measured the output of our laser using a power meter prior to each experiment. Briefly, the blue laser was turned on for at least 15 min to ensure power stabilization. Then the laser output power was calibrated by a compact power and energy meter console (PM100D, Thorlabs) both at the beginning and the end of each recording session. The output fiber diameter was 1 mm. In order to keep consistency, the distance from the fiber tip to the power meter detector was 1 mm.

We directly put the laser beam next to a temperature sensor. The value was averaged by 10 different measurements. In 250 mW laser stimulation for anesthetized rats, the average surface temperature was 60.2 ± 2.3 (mean $\pm$ sem) degree Celsius. The temperature for 150 mW stimulation was 53.9 ± 3.1 degree Celsius. However, such temperature measurements only provide a rough estimate of the true temperature achieved by the skin and we cannot be absolutely certain about the quantitative changes in skin temperature after a laser stimulus.

2.1.5. Data collection—Before recording, animals were given a 30 min period to habituate to a plastic recording chamber ($38 \times 20 \times 25 \text{ cm}^3$). Tetrodes were lowered in steps of $120 \mu\text{m}$ before each session recording. Rats were connected to recording equipment (Open Ephys) via an RHD2132 amplifier board and Serial Peripheral Interface (SPI) cable (Intan Technologies). Signals were monitored and recorded from 32 low-noise amplifier channels at 30 kHz, band-passed filtered (300 Hz to 7.5 kHz) with subsequent thresholding and then off-line sorted by a commercial software (Offline Sorter, Plexon). The threshold was lower than the 3 SD peak heights line and optimized manually based on the signal-to-noise ratio (SNR). The features of three valley electrodes were used for spike sorting. Only clear spike clusters with good tetrode spike waveforms, ISI (inter spike interval), auto- and cross-correlograms were selected for analysis. Units with firing rate lower than 1 Hz were excluded. In the end, we obtained well-isolated single unit activities. Trials were aligned to the initiation of paw withdrawal to compute the PETH for each single unit.

In the current study, we selected 10 recording sessions from 7 rats (3 from S1 and 4 from ACC, 5 sessions from each area) for the illustration purpose (table 1). All sorted single units are classified into putative pyramidal neurons and interneurons based on their waveform features, with roughly 10% of units being putative interneurons. All units were included in decoding analysis.

2.1.6. Postmortem histology—After completion of recording experiments, electrical lesions were made using $12 \mu\text{A}$ current for 30 s. Anesthesia was induced with isoflurane before perfusion through the heart with $1 \times$ phosphate buffered saline (PBS) followed by 4% paraformaldehyde (PFA). The brains were removed, fixed in 4% PFA, and equilibrated in 30% sucrose. Frozen sections were taken at $20 \mu\text{m}$ prior to cresyl violet staining. Positions of electrodes were determined by lesion point or trace mark identified under a compound microscope.

2.2. State space analysis

Let $k = 1, \dots, T$ denote the discrete-time index of a univariate or multivariate time series of neuronal spike counts (with a predefined bin size). Let $\mathbf{y}_k = [y_{1,k}, \dots, y_{C,k}]^\top$ denote a C -dimensional neuronal population vector (where the superscript $\top$ denotes the vector or matrix transpose), with each element denoting the observed neuronal spike count. Note that pain is an abstract emotional experience, and that pain signals are hidden and evolve dynamically in time. We proposed a latent variable framework to link the pain stimulus to neural activity.

For simplicity but without loss of generality, we made the following statistical assumption: (i) the latent variable $\mathbf{z}_k \in \mathbb{R}^m$ ($m-1$) represents the hidden common input that drives the neuronal population firing; (ii) the latent variable $\mathbf{z}_k$ follows a Gaussian Markovian process; (iii) the spike counts of individual units are Poisson distributed, and the latent variables and the observed variables are coupled by a Poisson linear dynamical system (PLDS), as follows (Macke *et al* 2012, Buesing *et al* 2013):

$$\mathbf{z}_k = \mathbf{A}\mathbf{z}_{k-1} + \boldsymbol{\varepsilon}_k \quad (1)$$

$$\boldsymbol{\eta}_k = \mathbf{C}\mathbf{z}_k + \mathbf{d} \quad (2)$$

$$\mathbf{y}_k \sim \text{Poisson}(\exp(\boldsymbol{\eta}_k)\Delta) \quad (3)$$

where state equation (1) is a first-order vector autoregressive (AR) model, with a state transition matrix $\mathbf{A} \in \mathbb{R}^{m \times m}$, $\boldsymbol{\varepsilon}_k \in \mathcal{N}(\mathbf{0}, \mathbf{Q})$ and $\mathbf{z}_1 \in \mathcal{N}(\mathbf{0}, \mathbf{Q}_1)$ specify temporal priors on the latent process. The parameters $\mathbf{d}$, $\mathbf{C}$ or $\boldsymbol{\eta}_k$ are unconstrained, and $\mathbf{C}$ can be a full (C -by- m) matrix, which allows the possibility that the individual neuronal firing is influenced by the latent population dynamics; Δ denotes the bin size. The observation equation is a GLM that employs the exponential link function through $\boldsymbol{\eta}_k$ (Truccolo *et al* 2005). The complete likelihood consists of two parts: one is Gaussian (from the state equation), and the other is Poisson (from the measurement equation). In sum, the complete-data log-likelihood of the PLDS consists of two parts: one from the latent variables, and the other from observed data

$$\begin{aligned}
\log p(\mathbf{y}_{1:T}, \mathbf{z}_{1:T} | \Theta) &= \log p(\mathbf{z}_{1:T} | \Theta) + \log p(\mathbf{y}_{1:T} | \mathbf{z}_{1:T}, \Theta) \\
&= \sum_{k=2}^T \log p(\mathbf{z}_k | \mathbf{z}_{k-1}) + \sum_{k=1}^T \log p(\mathbf{y}_k | \mathbf{z}_k, \Theta) \\
&= \sum_{k=2}^T (\mathbf{z}_k - \mathbf{A}\mathbf{z}_{k-1})^\top \mathbf{Q}^{-1} (\mathbf{z}_k - \mathbf{A}\mathbf{z}_{k-1}) - \frac{m}{2} \log(2\pi) - \frac{1}{2} \log |Q| + \sum_{k=1}^T \sum_{c=1}^C (y_{c,k} \eta_{c,k} - \exp(\eta_{c,k}))
\end{aligned}
\tag{4}$$

where $\Theta = \{\mathbf{A}, \mathbf{C}, \mathbf{d}, \mathbf{Q}\}$ denotes all unknown parameters of the model. Note that when equation (3) is replaced by a Gaussian likelihood model (i.e. $\mathbf{y}_k \sim \mathcal{N}(\boldsymbol{\eta}_k, \mathbf{R})$, where $\boldsymbol{\eta}_k$ and $\mathbf{R}$ denote the mean and covariance of the measurement noise, respectively), the PLDS reduces to the linear dynamical system (LDS) as a special case.

The goal of state space analysis is to infer the latent variables and unknown parameters given all the observed data (Chen *et al* 2010, Chen 2013, 2015). To maximize the objective function (4), likelihood inference can be tackled by an iterative expectation-maximization (EM) algorithm. In the E-step, we compute the Gaussian smoothed posterior for the latent state $\mathbf{z}_t \sim \mathcal{N}(\hat{\mathbf{z}}_t, \mathbf{P}_t)$; in the M-step, we update the parameters using the most recent state estimate. The iteration continues until the likelihood value is reached to the local maximum. Because of the non-Gaussian likelihood, the E-step in the EM algorithm is intractable. Therefore, Gaussian approximation methods can be considered for the PLDS, such as Laplace approximation (termed PLDSL_{laplace}) or variational approximation (termed PLDS_{variational}) (Smith and Brown 2003, Buesing *et al* 2013, Khan *et al* 2013). Specifically, we employed an efficient, convex dual variational inference algorithm (Khan *et al* 2013), where the unknown parameters are initialized by a subspace method (Buesing *et al* 2013, Pfau *et al* 2014).

Let $\mathbf{z} = [\mathbf{z}_1, \dots, \mathbf{z}_T]$ denote the joint latent state variable (size of mT -by-1) accumulated in time; the goal of variational inference is to maximize the lower bound of the marginal log-likelihood

$$\begin{aligned}
\log p(\mathbf{y}_{1:T}) &= \log \int \prod_{k=1}^T p(\mathbf{y}_k | \boldsymbol{\eta}_k) p(\mathbf{z}) d\mathbf{z} \\
&= \log \int q(\mathbf{z}) \frac{\prod_{k=1}^T p(\mathbf{y}_k | \boldsymbol{\eta}_k) p(\mathbf{z})}{q(\mathbf{z})} d\mathbf{z} \\
&\geq \mathbb{E}_{q(\mathbf{z})} \left[\log \frac{\prod_{k=1}^T p(\mathbf{y}_k | \boldsymbol{\eta}_k) p(\mathbf{z})}{q(\mathbf{z})} \right] \\
&= -\text{KL}(q(\mathbf{z}) \| p(\mathbf{z})) + \sum_{k=1}^T \mathbb{E}_{q(\mathbf{z})} [\log p(\mathbf{y}_k | \boldsymbol{\eta}_k)]
\end{aligned}
\tag{5}$$

where $\mathbb{E}_{q(\mathbf{z})}[\cdot]$ denotes mathematical expectation with respect to the variational distribution $q(\mathbf{z})$, $\boldsymbol{\eta}_k = [\eta_{1,k}, \dots, \eta_{C,k}]$ denotes the natural parameter, and $p(\mathbf{z}) = \mathcal{N}(\mathbf{0}, \mathbf{\Lambda}^{-1})$ denotes the prior with a tri-diagonal covariance structure of size $(mT \times mT)$

$$\Lambda = \begin{pmatrix} \mathbf{Q}_1^{-1} + \mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^\top & \mathbf{A}^\top\mathbf{Q}^{-1} & & & \\ \mathbf{Q}^{-1}\mathbf{A} & \mathbf{Q}_1^{-1} + \mathbf{A}\mathbf{Q}^{-1}\mathbf{A}^\top & \mathbf{A}^\top\mathbf{Q}^{-1} & & \\ & & \ddots & \ddots & \ddots \\ & & & \ddots & \ddots \end{pmatrix}$$

and $q(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{m}, \mathbf{V})$ denotes a variational Gaussian posterior with mean $\mathbf{m}$ and covariance $\mathbf{V}$. In variational inference, we aimed to maximize the lower bound of $\log p(\mathbf{y}_{1:T}|\Theta)$, which has the following form

$$\log p(\mathbf{y}_{1:T}|\Theta) \geq \mathcal{L}(\mathbf{m}, \mathbf{V}|\Theta) = \frac{1}{2}(\log|\mathbf{V}| - \text{tr}[\Lambda\mathbf{V}]\mathbf{m}^\top\Lambda\mathbf{m}) + \sum_{k=1}^T \mathbb{E}_{q(\mathbf{z})}[\log p(\mathbf{y}_k|\boldsymbol{\eta}_k)]$$

where $\text{tr}[\cdot]$ denotes the trace operator of a matrix.

In the M-step, we optimized $\{\mathbf{C}, \mathbf{d}\}$ using Newton's method to solve $\frac{\partial \mathcal{L}}{\partial \mathbf{C}} = 0$ and $\frac{\partial \mathcal{L}}{\partial \mathbf{d}} = 0$. In the E-step, the dual variational inference of $[\mathbf{m}, \mathbf{V}]$ uses a convex dual optimization method (Khan *et al* 2013) embedded with forward-backward Kalman smoothing. Specifically, let $\mathbf{y} = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_T]$ denote the augmented (CT -by-1) observation vector, let $\mathbf{d} = [\mathbf{d}_1, \mathbf{d}_2, \dots, \mathbf{d}_T]$ denote the augmented (CT -by-1) observation vector, and let $\mathbf{W}$ denote a CT -by- mT block-diagonal matrix,

$$\mathbf{W} = \begin{pmatrix} \mathbf{C} & & & \\ & \mathbf{C} & & \\ & & \ddots & \\ & & & \mathbf{C} \end{pmatrix}$$

Maximization of the variational lower bound can be formulated as a dual optimization problem (Khan *et al* 2013)

$$\min (\boldsymbol{\lambda} - \mathbf{y})^\top \mathbf{W} \boldsymbol{\Lambda}^{-1} \mathbf{W}^\top (\boldsymbol{\lambda} - \mathbf{y}) - \mathbf{d}^\top (\boldsymbol{\lambda} - \mathbf{y}) - \log|\boldsymbol{\Lambda} + \mathbf{W}^\top \text{diag}(\boldsymbol{\lambda}) \mathbf{W}| + \sum_{i=1}^{CT} (\lambda_i \log \lambda_i - \lambda_i) \text{ s.t. } \lambda_i > 0$$

Note that equation (6) is a convex function of parameter $\boldsymbol{\lambda}$. Let $\boldsymbol{\lambda}^*$ denote the optimal parameters derived from convex optimization, then the optimal variational parameters for $q(\mathbf{z}) = \mathcal{N}(\mathbf{z}|\mathbf{m}, \mathbf{V})$ are given by

$$\mathbf{m}^* = \boldsymbol{\Lambda}^{-1} \mathbf{W}^\top (\boldsymbol{\lambda}^* - \mathbf{y}) \quad (6)$$

$$V^* = (\Lambda + W^T \text{diag}(\Lambda^*) W)^{-1} \quad (7)$$

Instead of solving a direct mT -by- mT matrix inverse problem, calculation of the block-diagonal matrix V is equivalent to computing the smoothed posterior covariance in forward-backward Kalman smoothing for the LDS (Paninski *et al* 2009), which has a complexity of $\mathcal{O}(T(m^3 + m^2C + mC^2 + C^3))$. If $m = 1$, then the complexity reduces to $\mathcal{O}(TC + TC^2 + TC^3)$.

2.3. Statistical criterion for change

The Z-score is a standard metric to assess the change of neural activity in stimulus condition with respect to the baseline

$$Z\text{-score} = \frac{z - \text{mean of } z_{\text{baseline}}}{\text{SD of } z_{\text{baseline}}} \quad (8)$$

However, Z-score is often assessed for single unit activity based on multiple trials. In single-trial analysis, one can assess the equivalent Z-score based on the observed ensemble spike activity, where the mean and standard deviation are computed among the population within a specific single trial.

In the latent-variable framework, we propose to compute the Z-score from the estimated latent variable $\hat{\mathbf{z}}_{1:T}$. When the latent space is univariate (i.e. $m = 1$), the computation is identical to equation (8). Under the assumption that the Z-score is standard normally distributed, we convert it to the one-tailed p -value:

$$P(Z \times \text{score} > z) = 1 - P(Z\text{-score} \leq z) = 1 - \int_{-\infty}^z \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \quad (9)$$

When the latent state is multivariate normally distributed (i.e. $m > 1$), from the the estimated state posterior distribution $\mathcal{N}(\hat{\mathbf{z}}_{1:T}, \mathbf{P}_{1:T})$, we can similarly derive the p -value by using the following multivariate integration

$$P\text{-value} = 1 - \int_m^{\hat{\mathbf{z}}} \frac{1}{(2\pi)^{m/2} |\det(\Sigma)|^{1/2}} e^{-\frac{1}{2}(\mathbf{u} - \boldsymbol{\mu})^T \Sigma^{-1} (\mathbf{u} - \boldsymbol{\mu})} d\mathbf{u} \quad (10)$$

where $\boldsymbol{\mu}$ and Σ denote the time-averaged mean and covariance of the latent variable in the baseline, respectively. In MATLAB (The MathWorks, Natick, MA) implementation, we computed the one-sided p -value as $1 - \text{normcdf}(Z, 0, 1)$, where normcdf denotes a normal cumulative distribution function, or as $1 - \text{mvncdf}(\mathbf{z}, \boldsymbol{\mu}, \Sigma)$. If the latent variable is non-Gaussian, we can replace the Gaussian with a long-tailed student t -distribution to derive the

p -value. In MATLAB command, it reads as $1 - \text{tcdf}(Z, \nu)$, where ν denotes the degrees of freedom. The t -distribution becomes closer to the normal distribution as ν increases.

The criterion of Z-score change is determined by a pre-described threshold ρ , which is also related to the desired p -value for reaching statistical significance. For instance, using the significance criterion with one-sided p -value 0.05, it is concluded that when Z-score > 1.65 or Z-score < -1.65 (where 1.65 is the critical value for reaching significance under the normality assumption), the stimulus-induced population activity increases or decreases its firing significantly from its baseline condition, respectively. When the threshold ρ is increased to 2.33, 2.58, 3.09, 3.32 and 3.52, the corresponding p -value will be 0.01, 0.005, 0.001, 0.0005, and 0.0001, respectively.

2.4. Online prediction and evaluation on testing data

Upon identifying the model parameter $\Theta = \{\mathbf{A}, \mathbf{C}, \mathbf{d}, \mathbf{Q}\}$, we can run an online forward filter to recursively infer the state variable

$$\hat{\mathbf{z}}_{k|k-1} = \mathbf{A}\hat{\mathbf{z}}_{k-1|k-1} \quad (11)$$

$$\mathbf{P}_{k|k-1} = \mathbf{A}\mathbf{P}_{k-1|k-1}\mathbf{A}^\top + \mathbf{Q} \quad (12)$$

$$\hat{\mathbf{y}}_{k|k-1} = \exp(\mathbf{C}\hat{\mathbf{z}}_{k|k-1} + \mathbf{d})\Delta \quad (13)$$

$$\mathbf{P}_{k|k}^{-1} = \mathbf{P}_{k|k-1}^{-1} + \mathbf{C}^\top \text{diag}(\hat{\mathbf{y}}_{k|k-1})\mathbf{C} \quad (14)$$

$$\hat{\mathbf{z}}_{k|k} = \hat{\mathbf{z}}_{k|k-1} + \mathbf{P}_{k|k}\mathbf{C}^\top(\mathbf{y}_k - \hat{\mathbf{y}}_{k|k-1}) \quad (15)$$

where $\mathbf{P}_{k|k-1}$ and $\mathbf{P}_{k|k}$ denote the predicted and filtered state covariance matrices, respectively.

In addition, we can evaluate the log marginal likelihood of new data $\mathbf{y}$ by integrating out all latent variables

$$\begin{aligned}
\log p(\tilde{\mathbf{y}}_{1:T} | \Theta) &= \log \int p(\tilde{\mathbf{y}}_{1:T}, \mathbf{z}_{1:T} | \Theta) d\mathbf{z}_{1:T} \\
&= \sum_{k=1}^T \log p(\tilde{\mathbf{y}}_k | \tilde{\mathbf{y}}_{1:k-1}) \\
&= \sum_{k=1}^T \log \int p(\tilde{\mathbf{y}}_k | \mathbf{z}_k) p(\mathbf{z}_k | \tilde{\mathbf{y}}_{1:k-1}) d\mathbf{z}_k
\end{aligned} \tag{16}$$

For the PLDS, computation of equation (16) can be done either by variational approximation, or by Laplace approximation (Shimazaki *et al* 2012), or even by Monte Carlo approximation as follows

$$\log p(\tilde{\mathbf{y}}_{1:T} | \Theta) \approx \frac{1}{N_p} \sum_{i=1}^{N_p} \sum_{k=1}^T \log p(\tilde{\mathbf{y}}_k | \mathbf{z}_k^{(i)}) \tag{17}$$

where $\mathbf{z}_k^{(i)} \sim \mathcal{N}(\hat{\mathbf{z}}_{k|k-1}, \mathbf{P}_{k|k-1})$ are independent and identically distributed (i.i.d.) samples drawn from the approximate Gaussian posterior $p(\mathbf{z}_k | \tilde{\mathbf{y}}_{1:k-1})$. The predictive likelihood can be used for model selection—for instance, for choosing the dimensionality of latent variable (see table 2 for a demonstration). In this illustrated example, the predictive likelihood increased with dimensionality m for PLDSL_{aplace}, but reached a local maximum $m = 2$ for PLDS_{variational}. PLDS_{variational} had a slightly better performance than PLDSL_{aplace} for lower dimensionality ($m < 3$). However, in cross-validation experiments, we did not see significant differences in performance between $m = 1$ and $m = 2$. In addition, considering the fact of significantly increased computational cost in inference and Z-score computation, we chose $m = 1$ for the latent state dimensionality.

2.5. Special case of the model

In the case of $C = 1$ (i.e. single neuron observation), the PLDS model (equations 1 through 3) reduces to a simplified form

$$z_k = a z_{k-1} + \varepsilon_k \tag{18}$$

$$y_k \sim \text{Poisson}(\exp(c z_k + d) \Delta) \tag{19}$$

where $0 < |a| < 1$, $c \in \mathbb{R}$ and $d \in \mathbb{R}$. Because of the scale ambiguity, without loss of generality we can set $c = 1$ and $d = 0$. In this case, identification of the latent variable $\{z_k\}$ is equivalent to estimating the smoothed time-varying Poisson firing rate $\lambda_k = \exp(z_k)$ of the single neuron (Smith and Brown 2010). Here, $\{y_k\}$ is a doubly-stochastic Markov-driven Poisson process, with variance $\text{Var}[y_k] = \mathbb{E}[y_k] + \text{Var}[\exp(z_k)]$, and the random variable $\exp(z_k)$ is lognormally distributed.

In the context of decoding pain signals, if the single neuron is a pain-modulated (either positively or negatively modulated) unit, the smoothed Poisson firing rate of that pain-modulated neuron will determine the presence or absence of a pain stimulus. In this special case, we can establish the link between the Z-score and the SNR. Let

$SNR = \frac{\sigma_{\text{signal}}^2}{\sigma_{\text{noise}}^2} = \frac{\sigma_{\text{signal}}^2}{\sigma_{\text{baseline}}^2}$ and $SNR - 1 = \frac{\sigma_{\text{signal}}^2 - \sigma_{\text{baseline}}^2}{\sigma_{\text{baseline}}^2} \approx \frac{\lambda_{\text{signal}} - \lambda_{\text{baseline}}}{\lambda_{\text{baseline}}}$, where λ_{signal} and $\lambda_{\text{baseline}}$ denote the respective marginal firing rate parameters in signal and baseline conditions (Poisson mean equal to Poisson variance). Comparing SNR-1 and Z-score (equation (8)), it appears that SNR - 1 computes the relative score of marginal firing rate, whereas Z-score computes the relative score of log firing rate ($z_k = \log \lambda_k$).

3. Results

3.1. Acute thermal pain behavior

Earlier studies have used CO₂ lasers to produce thermal pain stimuli (Kuo and Yen 2005, Zhang *et al* 2011, Wu *et al* 2016). Here, we chose to use a blue laser for the following reasons. Since a blue laser does not produce heat as intense as a CO₂ laser, we can adjust the power output of the laser to create different thermal stimuli. By varying the thermal intensity, we can use paw withdrawal as an index for the verification of thermal pain. A second advantage of the blue laser is that due to the relatively lower level of heat it can generate, we can repeat the application of the stimulus at the same spot in the paw without worrying about hypersensitivity or tissue injury. Application at the same anatomic location is particularly important for our decoding exercise in the S1 region.

In the experiment, we delivered laser stimuli to the plantar surface of the hind paw (contralateral to the recording site) of freely behaving rats. It was natural to expect subject variability in animals' pain behavior. We used paw withdrawal as the behavioral readout for pain, as it is the standard practice in rodent studies. Since associated motion may lead to artifacts that mask the true pain signal, we used video to verify the onset of withdrawal response to pain. Summary of the behavioral and physiological datasets used in the current study is shown in table 1. It is important to note that paw withdrawals occur nearly 100% in all laser stimulations with intensity 250 mW. In contrast, for the 50 mW laser stimulation, around 90% of trials did not elicit paw withdrawals (figure 2(B)). The paw withdrawal behavior also occurs in response to alternative acute mechanical pain stimuli, such as pin prick. As expected, the laser-to-withdrawal latency decreased with increasing laser intensity (figure 2(C)).

Overall, we observe a clear temporal association between the laser (or pinprick) stimulus and paw withdrawal, leading us to believe that such withdrawal events are caused by acute pain. Of course, we cannot completely preclude the possibility that occasionally a paw withdrawal event is due to locomotion or other cues, but this is a standard limitation for most pain behavior testing using paw withdrawal as an index.

3.2. Single units of S1 and ACC encode pain signals

We used tetrode arrays to record *in vivo* ensemble spike activity from the rat S1 and ACC. We did not map the receptive fields of S1 neurons. Instead, we have chosen the location of our S1 electrodes based on previous reports on S1 receptive fields for the paw (Armstrong-James and George 1988, Haupt *et al* 2004). On the other hand, there is less evidence in the literature on a somatotopic map in the ACC, therefore we inserted tetrodes based on the literature on the general area within the ACC that is responsible for pain processing. Upon off-line spike sorting, we obtained well-isolated single-unit activity (SUA). The spikes at each trial were aligned with paw withdrawal, from which we could obtain the respective peri-event time histograms (PETH) in response to the stimulus or behavioral response.

To identify a ‘pain-modulated’ single unit, we used the following statistical criteria in our current decoding analysis: (i) Based on the spike rate's confidence intervals, if the lower bound of the post-event firing rate is at least three standard deviation above the pre-event (baseline) firing rate, we will conclude there is a significant firing rate increase and the unit is a ‘positive responder’. We used a (mean + 3SD) criterion for assigning statistical significance, where the baseline is defined as the 5-s period before the laser stimulus. (ii) Alternatively, if the upper bound of the post-event firing rate is below the lower bound of the baseline firing rate, then the unit is a ‘negative responder’. We used a (mean – 3SD) criterion for assigning statistical significance. If the unit is neither a positive responder nor a negative responder, then it is assigned as a non-responder. (iii) The first time point that the spike rate meets the first or second criterion is used to define the neuronal response latency.

The illustration, Figure 2 shows the raster plots and PETHs of representative ‘positive responder’ and ‘negative responder’ from the rat ACC and S1. Alternatively, one can align the raster with the laser onset to obtain a different representation of neural responses. It is important to note that the identification of positive or negative responder remain robust regardless of the alignment of paw withdrawal or laser onset. Notably, for pain-modulated units, the SNR is much higher in trial-averaged PETHs than in single-trial rasters. The next question is: can we reliably detect the pain signal from these pain-modulated units on a single-trial basis? Note that since we emphasize single-trial decoding, we do not use PETH for the decoding purpose.

3.3. State space modeling

The motivation of state space modeling is to develop a data-driven statistical model to characterize the change in the network in response to pain stimuli. Since neurons receive a common input, their responses are often temporally coordinated. If we treat a pain-modulated neuronal response as a signal, and a pain-unmodulated neuronal response as noise, then the SNR of the neuronal population is directly relevant to the detection. The higher the SNR, the easier is the detection task. In general, the creation of a whole (population) is greater than the simple sum of its parts in terms of detection power and SNR.

Based on the empirical observations in neural encoding, we find a proxy for the acute pain signal that drives the observed neuronal population spike activity. We propose a latent SSM, also known as PLDS, to capture the temporal dynamics of the observed multivariate

neuronal spike count $\mathbf{y}_k$ from C neurons and the latent acute pain signal $\mathbf{z}_k$, from which we further infer the pain onset. In principle, we could use model selection criteria to select the dimensionality of m for the latent state. In practice, however, we did not find any performance difference in detecting acute thermal pain signals between using $m = 1$ or $m > 1$. Therefore, in the results reported below, we have used $m = 1$ in all visualizations.

It is also noteworthy that the inferred univariate latent variables are different from the mean of the observed neuronal responses, namely $z_k \neq \sum_{c=1}^C y_{c,k}$, for several important reasons: (1) the presence of noise; (2) nonlinearity because of Poisson firing and log-link function; (3) smoothing; and (4) baseline offset. Our model-based inference procedure employs a forward-backward smoothing operation to estimate the optimal latent state and unknown parameters. In addition, when $m = 1$, the entries of (vector) parameter $\mathbf{C}$ specify the positive or negative contributions of the latent variable to each neuronal response.

3.4. Single-trial population decoding analysis

The goal of population decoding is to identify or detect the onset of an acute thermal pain signal in neural codes. Unlike traditional neural coding paradigms where the decoded variables are measurable (such as the visual stimulus, movement kinematics, or animal's position), the pain signal is hidden. Based on the observation of neural encoding, we can find a proxy for the acute pain signal that drives the observed neuronal population spike activity. Specifically, the decoding was formulated as a change detection problem, whereas the change is defined related to the assumed baseline condition. In the single-trial examples, we defined the baseline as 4 s prior to the laser onset as the baseline period. However, our results reported below were also robust to different lengths of the baseline period.

First, we applied our methods to analyze ACC and S1 ensemble spike data in an individual single-trial setting. All sorted single units (putative pyramidal neurons and interneurons) were included in decoding analyses. Figure 3 present representative S1 and ACC ensemble examples of single-trial decoding analyses, both under 150 mW laser stimulation. The S1 and ACC populations had simultaneous recordings of 12 and 7 single units, respectively. There was strong variability of population firing between single trials, in terms of the baseline firing, paw withdrawal latency, and neuronal response patterns. Notably, this S1 population example mostly consisted of positive responders; whereas this ACC population example consisted of both positive and negative responders. The opposite raster examples are shown in figure 4.

As a comparison, we also computed the ‘empirical’ Z-score from the raw multi-unit spike activity (i.e. the total spike count $\sum_{c=1}^C y_{c,k}$). The empirical Z-score was computed by the total spike count minus the averaged total spike count during the baseline period, and further normalized by the standard deviation. It is worth pointing out some interesting phenomena in figure 3. First, the model-derived Z-score trace (blue) was much smoother than the empirical Z-score trace (red). Second, when the population consists of only positive responder units (such as figure 3(A)), the two Z-score traces followed a similar trend. However, when the population consists both positive and negative responder units (such as figure 3(D)), the empirical Z-score metric failed because the contributions of spike activity

from positive and negative units canceled out. In contrast, the model-based approach was still able to detect the onset of the pain signal.

An alternative strategy is to treat each neuron independently and compute the empirical Z-score (absolute value) of each unit. However, in order to reach a detection decision, we need to develop a pooling or voting strategy. Since we don't have any prior knowledge of those neurons, we treat every neuron equally and independently. Specifically, employing different voting rules (e.g. arithmetic mean, geometric mean, maximum, or minimum operation) would result in different detection outcomes. More importantly, regardless of the voting system being used, the detection threshold trajectory would be 'oscillatory' (i.e. not smooth) due to lack of smoothing. In addition, treating neurons independently would discard important temporal information coordinated in the neuronal responses. For the purpose of illustration, using the same data in figure 3(D), we present three different detection outcomes in figure 5 by applying the 'mean' and 'maximum' voting strategies on the empirical spike activity. As compared to figure 3(E), all of these absolute Z-score traces are noisy (due to lack of smoothing) and are not always able to reach a consensus (except around 7.3 s). As expected, the 'maximum' strategy would be sensitive to outliers and create many false positives. Therefore, it is hard to design an optimal voting strategy *a priori* based on individual spike count series. In contrast, our state space modeling provides a principled approach for temporal smoothing and integration of population spike activities.

As illustrated in these two representative examples, our approach successfully detected the 'neuronal threshold for acute pain' from both S1 and ACC populations separately. The statistical significance was assessed by the p -value. Using a significance criterion of $p = 0.05$, we could detect pain signals in 38 out of 40 trials in the rat1-S1-1 population example, and in 16 out of 16 trials in the rat4-ACC-1 population example. Changing the significance criterion affected the detection results. In the rat1-S1-1 example, the ratios of pain detection events were 32/40, 29/40, 22/40, 15/40 and 11/40 for the significance criterion of p -value 0.01, 0.005, 0.001, 0.0005 and 0.0001, respectively. Whereas for the rat6-ACC-1 example, the ratios of pain detection events were 13/16, 12/16, 11/16, 11/16 and 9/16, respectively.

Two additional single-trial decoding examples for S1 and ACC ensembles are shown in figure 4. The additional example of S1 ensemble consists of both positive and negative responders, whereas the additional ACC ensemble consists of both positive responders and non-responders.

From detection analyses, we defined the network response latency as the first time point that the Z-score reached the statistical significance level relative to the onset of paw withdrawal in response to pain. Our analyses showed that the network response latency was negative in nearly half of the single trials for both S1 (e.g. figure 4(B)) and ACC (e.g. figure 5, right panels). This suggests that in principle we may detect the pain signal in neural codes before the actual spinal reflexive pain behavior (paw withdrawal).

3.5. Sequential decoding analysis

Next, we applied the single-trial decoding analysis in a continuous or sequential manner. In other words, we did not separate individual single trials and defined the baseline as the first

4 s before the very first trial. The purpose of this analysis is to test the robustness of the decoding approach, since the Z-score is computed with respect to the baseline and we expect that there is variability in the spontaneous (baseline) activity in time. There are two computational options to be considered. The first option is to run the inference and smoothing over the entire time series (See figure 6 for an illustration), whereas the second option is to only run inference and smoothing over a fixed-length time series (training set), and then use the model to infer the latent variable from the remaining time series (testing set). The first one is more computationally intensive when the duration of the time series is long. The second option is computationally faster, but its performance depends on the stationarity of the statistical model (including the baseline). If there is a model mismatch between the training and testing sets, then the performance will degrade. In the current analysis, we restricted ourselves to the first option.

3.6. Multi-unit activity (MUA) decoding

Thus far we have restricted our analyses to the sorted SUA. However, spike sorting is a time-consuming and error-prone process (Lewicki 1998). In order to apply our approach to the clinical application of deciphering the onset of pain, it is, therefore, more appealing to perform population-decoding analysis based on unsorted MUA (Chen *et al* 2012, Fraser *et al* 2009, Kloosterman *et al* 2014, Ventura 2008). Under the Poisson firing assumption, since the sum of Poisson variables is Poisson distributed, the MUA at each tetrode is also Poisson distributed. In general, the detection sensitivity and specificity would degrade because of two reasons: (1) lower SNR; and (2) lack of neuronal specificity (firing rates of individual units either increase or decrease at a given time point, summation of them would cancel out the effect).

We replaced the ‘unit’ with ‘tetrode’ (which consists of MUA) and tested the decoding accuracy using the same probabilistic approach. Surprisingly in many examples, our proposed approach performed quite well in detecting pain signals from unsorted MUA (see figure 7 for an illustration). In general, due to the lack of unit selectivity, we expect that the performance would be less optimal (figures 7 and 8, see also table 3). Specifically, comparing figure 6 with figure 7, we observed that the sequential MUA decoding had a significantly degraded accuracy, with increases in both false positives and false negatives. This could be ascribed to the nonstationarity of the data or the change in MUA baseline, since the SUA and MUA decoding had comparably accuracy in the (non-sequential) single-trial decoding setup (table 4, Dataset rat2-S1-2).

3.7. Sensitivity and specificity

It is important to assess the robustness of our proposed pain detection method, especially the detection specificity (i.e. true negative rate) and sensitivity (i.e. true positive rate). In other words, we need to determine the type-I error (false positive) and type-II error (false negative) in relation to the detection threshold. To do so, we repeated the decoding analysis for all single trials under the same laser intensity, and recruited an equal number of negative control trials in the absence of a pain stimulus (i.e. the non-overlapping between-trials period). We computed the error by varying the detection threshold and further computed the detection sensitivity and specificity (Macmillan and Creelman 2004):

$$\text{Sensitivity} = \frac{\# \text{true positives}}{\# \text{condition positives}},$$

$$\text{Specificity} = \frac{\# \text{true negatives}}{\# \text{condition negatives}}$$

where the number of condition positives is equal to the sum of the number of true positives and the number of false negatives; the number of condition negatives is equal to the sum of the number of true negatives and the number of false positives. One quantitative way to visualize these two metrics is the ROC (receiver operating characteristic) curve, with the x -axis being (1-Specificity) and y -axis being Sensitivity. By systematically varying the significance threshold of detection, we obtained corresponding statistics for sensitivity and specificity and obtained multiple sample points in the ROC curve. The ROC plot shows the tradeoff of sensitivity and specificity at different thresholds, and the area under ROC (AUROC) indicates the overall detection performance (1 being perfect, 0.5 being a chance level).

As an illustration, Figure 9 shows the ROC curves derived from decoding results of rat-S1-1 and rat4-ACC-1 population examples using sorted ensemble spike activity. The AUROC values were 0.983 and 0.963 for the S1 and ACC examples, respectively, suggesting an excellent detection/discrimination performance. In these plots, we recruited equal numbers of negative control trials in the absence of a pain stimulus (during between-trial periods). Table 4 lists the AUROC statistics from the selected datasets. As seen from the table, the overall AUROC statistics from SUA decoding were excellent for both S1 (mean $\pm$ SD: 0.949 $\pm$ 0.06) and ACC (0.922 $\pm$ 0.04). In contrast, the AUROC statistic from MUA decoding was generally worse for ACC (0.866 $\pm$ 0.02) than S1 (0.892 $\pm$ 0.09, the majority was greater than 0.9 except for the dataset rat4-S1-4 with only one tetrode).

For all single trials, we computed the true positive rate using either sorted SUA or unsorted MUA. The results are summarized in table 4. We found that in the case of SUA decoding, for either $p = 0.05$ or $p = 0.001$, the false negative rate (1 – true positive rates) of S1 was lower than that of ACC; whereas in the case of MUA decoding, the S1 and ACC were relatively comparable in true positives.

3.8. Percentage of modulated units correlates with the Z-score amplitude

Pain intensity is related to the property of the noxious stimulus, and previous studies have correlated stimulus intensity with relative changes in neuronal firing rates (Zhang *et al* 2011). Thus, it is possible that neuronal firing rates can be correlated with pain intensity. In our experiments, we have observed variations in neural firing rates within a single session. Here in our computational analysis, we used the peak value of (Z-score minus confidence interval) as an indicator of qualitative pain intensity. Generally, a higher Z-score peak magnitude indicates a larger increase in the population firing rate, implying higher pain intensity. In addition, the percentage of modulated units may serve as another indicator for

pain intensity. The information is captured by matrix $\mathbf{C}$ of the observation equation (equation 2). When $m = 1$, we can monitor the amplitude of the product ($\mathbf{C} \times \mathbf{Z}$ -score), each element of which indicates the relative contribution from each neuron. In principle, a high percentage of modulated units implies that most elements (in absolute value) are greater than 1, whereas a unmodulated unit has a small (absolute) value. Based upon single-trial analyses in rat1-S1-1 and rat4-ACC-1 datasets, we found that the maximum Z-score was positively correlated with the percentage of modulated units ($R = 0.38$ and $p = 0.015$ for rat1-S1-1, $R = 0.638$ and $p = 0.008$ for rat4-ACC-1, Spearman's rank correlation), using a high significance detection criterion ($p = 0.001$). These results suggest that we can qualitatively decode the pain intensity. Two additional examples (Datasets rat3-S1-3 and rat6-ACC-3) are illustrated in figure 10. As shown in the figure, there was a wide-range of between-trial variability in the maximum Z-score and ratio of modulated units.

3.9. SNR comparison of population code and single neuron

How does the population code compare with a single neuron in reading out acute pain signals from a decoding perspective? The answer to this question pertains to the SNR in single trials. In the Methods section, we have shown that in the special case when the PLDS has only a single neuron, estimating the latent variable is equivalent to estimating a smoothed time-varying Poisson firing rate of that given neuron. The SNR of a single neuron is related to the relative firing rate change in the pain-induced response from the baseline. Imagine that we can find a 'best-tuned' pain-modulated neuron, adding additional neurons may not necessarily further improve the SNR or decoding accuracy. On the other hand, if the single neuron is not perfectly modulated with pain stimuli, adding more neurons is likely to improve the SNR and the decoding performance. Therefore, what matters most is not the total number of recorded neurons, rather the percentage of pain-modulated neurons in recordings.

In unsorted MUA, the spike activity from each tetrode may contain multiple units and additional noise. On average, the SNR of single tetrode is lower than the SNR of single unit. Therefore, increasing the number of recorded neurons has a potential to increase the population SNR in MUA decoding.

As expected, the detection performance depends on the number of recorded neurons. To investigate the impact of the number of neurons on the performance, we used one representative population example (Dataset rat1-S1-1, 12 neurons) and randomly selected a subset of the neurons to conduct the decoding analysis. Specifically, we varied the population percentage as 8.3%, 25%, 50% and 75% (i.e. the total number of neurons used are 1, 3, 6 and 9, respectively) and repeated the random selection 50 times. The result of detection accuracy (true positive rate, significance threshold $p = 0.05$) is summarized in figure 11. Naturally, decoding accuracy increased as more neurons were used, with the degree of improvement depending on the total number of neurons and the responsiveness of individual neurons in each recording. In the extreme case of using a single neuron, the performance is around the chance level.

3.10. Robustness to nonstationarity

Our modeling and inference assume model stationarity among the temporal data being analyzed (see figure 12 for illustrations in two baseline periods). As an example to test this assumption, we used the the first trial of 10 s S1 recording (Dataset rat1-S1-1) to estimate the model parameters, and then fixed those parameters to infer unknown latent state variables and Z-scores (using online forward filtering) in the consecutive trials. As a comparison, we plot the Z-scores computed from the standard EM inference and online filtering algorithm (equations 11 through 15). In this specific example (figure 13), the detection results are very close. Notably, the computational speed in the latter case is nearly 30 times faster.

In practice, this assumption is sometimes violated due to intrinsic non-stationary spontaneous firing activity, especially in MUA recordings where units could emerge or disappear in time (e.g. figure 7). Additionally, in freely behaving animals, the S1 and ACC ensemble activities are also modulated by other factors, such as the tactile stimulus, reward-related action and emotion. The nonstationarity of the baseline activity poses a great challenge to reliably detect the onset of acute pain signals in a sequential context.

One research direction under consideration is to introduce online adaptation to the model in time, so that the baseline and model parameters are updated. Another solution is to integrate useful multi-source physiological signals in addition to neuronal ensemble spike activity, such as the heart rate, blood pressure, and local field potentials (LFPs). Extensive investigations of these directions are beyond the scope of the current paper.

4. Discussion

4.1. Population codes for acute pain

Pain is a complex perceptual process that involves multiple areas of central and peripheral nervous systems (Perl 2007). Although several acute pain studies have been recently dedicated to the neural spike activity in rodent S1 and ACC circuits (Kuo and Yen 2005, Zhang *et al* 2011), a complete understanding of neural mechanisms for acute pain remains elusive (Sorkin and Wallace 1999). In the animal model of acute pain, several challenges are noticeable: first, unlike traditional sensory (e.g. visual and auditory) stimuli, it is difficult to standardize pain stimuli, especially in freely behaving animals. Second, since many recorded neuronal responses are not pain-specific, the coding strategy for pain can be very complex and remains unclear to date (Prescott *et al* 2014). Third, in addition to the physical properties of pain stimuli, pain signals are modulated by other inputs, such as anxiety, fear, anticipation, adaptation, motoric/autonomic shifts, and motivational/attentional effects (Coghill *et al* 1999, Craig 2002, 2003, Apkarian *et al* 2011). How to integrate this information to derive pain signals or pain experiences remain unsolved.

In the current study, we have focused on a simple acute pain model using noxious thermal stimuli. As the first step, we have focused on the population codes in S1 and ACC. However, we expect that neural mechanisms of acute pain involves many other important neural circuits.

Decoding acute pain signals is viewed as a latent-variable change-point detection problem. From a decoding perspective, the SNR of population codes in single trials is lower. Extracting pain signals requires proper temporal smoothing and baseline comparison. To achieve this goal, we propose a PLDS model to detect the change in the latent variable that drives the population spike activity, and we use a variational EM algorithm to identify the unknown state and model parameters. The latent variable can be univariate or multivariate, but using a univariate latent variable is not only computationally efficient (without compromising the detection performance), but also easier for interpretation of the Z-score. To investigate the sensitivity and specificity of the proposed approach, we apply the methods to negative controls. In general, the S1 yields a higher false positive rate compared to the ACC, and this could be due to the fact that S1 is more responsive to sensory stimuli including non-pain stimuli.

4.2. Insights

Our data-driven decoding approach for detecting acute thermal pain signals provide further scientific insights and new strategies. First, statistical modeling of S1 or ACC populations allow us to capture the contribution of the latent variable to the firing of each neuron, and to reveal the underlying latent dynamics and noise statistics. Second, in terms of detection specificity and sensitivity, under the same statistical criterion, S1 has better sensitivity and less specificity, whereas ACC has better specificity and less sensitivity. Therefore, if we have the access of simultaneous S1 and ACC recordings, we can design independent statistical significance criteria for these two neuronal ensembles (e.g. using a stricter significance criterion for S1), in order to derive an optimal coding strategy.

For instance, since the reliability of S1 and ACC population-decoding differs, we can design an optimal strategy to weight the contributions from S1 and ACC ensembles. Let P_{S1} and P_{ACC} denote the posterior distributions of detecting pain signals computed from S1 and ACC, respectively; and let $0 < \pi_{S1} < 1$ and $0 < \pi_{ACC} < 1$ denote their respective prior probabilities (where $\pi_{S1} + \pi_{ACC} = 1$); then the ultimate posterior distribution of detecting the pain signal is governed by

$$P = \pi_{S1} P_{S1} + \pi_{ACC} P_{ACC}.$$

Note that the significance criterion for detection can be different in these two regions. We will test the optimal ratio π_{S1}/π_{ACC} in the future.

4.3. Computational speed

The maximum likelihood inference for the PLDS model consisted of state and parameter estimation, which was fulfilled by an iterative EM algorithm. A naive implementation of the variational EM algorithm has computational complexity of order $(x1D4AA)(Lm^3T)$, where L denotes the number of EM iterations until convergence. In our non-optimized MATLAB (MathWorks, Natick, MA) implementation, for a single-trial analysis of 12 units with 200 temporal bins, the CPU time for EM convergence was 10–20 s on a MacPro computer (2.7 GHz 12-Core, 64 GB RAM). However, it shall be noted that most computation overhead

involved iterative estimation for parameters. If the parameters are identified *a priori* from the baseline period, the time for inferring latent state $\mathbf{z}_t$ (i.e. one-pass of the E-step) is much smaller (around 2 s).

To improve the inference speed for the PLDS model, we considered two strategies. First, we replaced the variational inference with Laplace inference in the E-step of EM algorithm. In this case, the computation speeds up dramatically. Second, we simplified the PLDS to a LDS by approximating Poisson observations with half-rectified Gaussian observations. Because of the Gaussian likelihood used in the observation equation, the E-step of the EM algorithm ran much faster, and the algorithm converged within a few hundred milliseconds. Note that at each EM iteration, the computational bottleneck is the variational E-step. When $m = 1$, the complexity will only scale linearly with C . In online filtering setup, the complexity will be further reduced.

In terms of the trade-off between computational complexity and decoding performance, we found that the LDS has the least computational complexity but worst decoding accuracy (for both false positives and false negatives), the PLDS_{variational} has the highest computational complexity and the best decoding accuracy, and the PLDS_{Laplace} has modest computational complexity (with convergence speed similar to LDS for a small T) and good performance comparable to PLDS_{variational}. Therefore, PLDS_{Laplace} may be a good candidate used for real-time diagnostic applications.

4.4. Unsupervised versus supervised strategies

Our proposed method for detecting acute pain signals is unsupervised in that it does not use any training labeled data and we do not explicitly use any knowledge of laser onset or withdrawal information. In contrast, we may also use supervised learning algorithms, such as the support vector machine, to classify acute pain and non-pain signals based on spatio-temporal patterns (that may include joint observations of spikes and field potentials) (Chen *et al* 2007, Sowards *et al* 2012). Compared to the unsupervised approach, a supervised approach requires collecting a large number of trials or training samples with labels, for both positive and negative samples. However, we may encounter a sample imbalance between the positive and negative labels. Furthermore, the classifier needs to be retrained for different sessions (depending on different number of neurons), whereas our proposed unsupervised approach is relatively independent on the number of neurons. In addition, supervised approaches often require a large number of samples to cross-validate the free model parameters (such as the regularization parameter, margin parameter, etc). In practice, we can adapt different strategies according to the need. For instance, we can use an unsupervised learning approach to detect the pain signals, and then use a supervised learning approach to detect the pain level associated with different laser intensities. Ultimately, many supervised approaches can be combined with our unsupervised approach to establish a more robust method for detecting pain signals.

4.5. Spikes versus LFPs

In addition to spike activity, another important signal from extracellular measurements is LFP. LFP represents the aggregate subthreshold activity of a local population of neurons in a

spatially localized area near the recording electrode and can be viewed as the input information in that area. Although LFPs carry certain information of pain, such as laser-evoked potential (LEP), they lack the cellular resolution to examine the detailed representation. Due to low SNR, single-trial analysis of LFP is very challenging. In addition, there are potential artifacts corrupting the LFP signals when using freely behaving animals, which makes the analysis even more challenging. Finally, local channels of LFP tend to be highly correlated, therefore there is no sufficient spatial sampling diversity in LFP as seen in neuronal spike activity. However, denoised LFP signals can be used for post-hoc confirmation of pain detection. In addition, from the encoding perspective, it is interesting to examine the spike-LFP dependence, such as spike-field coherence (SFC).

5. Conclusion

In conclusion, we developed a SSM-based approach for detecting acute thermal pain signals based on the ensemble spike activity recorded from the rat S1 and ACC. Our investigations have showed efficacy and robustness of our approach in various conditions. These results point out a promising direction for detecting acute pain signals in real time. Although we use the acute thermal pain model here, it would be interesting to extend the decoding analysis to other pain models (Kim *et al* 1997, Boyce-Rustay *et al* 2010, Jaggi *et al* 2011, Xu and Brennan 2011). For instance, we can use a rat model of chronic pain and apply acute thermal pain stimuli to the animal in order to compare its pain behavior or assess changes in neuronal responses associated with chronic pain. We can also use decoding analysis as an alternative behavior readout to compare the sensitivity or specificity of neuronal populations at different brain regions. Much remains to be done both experimentally and computationally to unravel the pain mechanisms and their links to behavior.

An intermediate goal is to adapt this decoding approach for a real-time brain-machine interface (BMI) system, where the onset of an acute pain signal will be used to trigger optogenetic control for pain modulation (Gu *et al* 2015, Lee *et al* 2015, Copits *et al* 2016, Iyer *et al* 2016). A long-term research goal is to employ minimally invasive ECoG or noninvasive EEG recordings in patients to monitor their brain activity, from which we can make useful diagnoses or quantification of the pain. The current study that tests the feasibility of decoding acute pain in a rodent model represents an initial step towards our ultimate research goal. There have been notable successes in related fields such as population decoding in the primary motor cortex, where considerable progress has been made to translate initial understanding of rodents or nonhuman primates to humans.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the US NSF-CRCNS grant IIS-130764 (ZC), NIH-CRCNS grant R01-NS100065 (ZC, JW), and NIH grants K08-GM1026911 and R01-GM115384 (JW). We thank Dr Lars Buesing for sharing the MATLAB code for implementing variational inference. For result reproducibility, we provide a MATLAB demo code and two representative datasets in the supplementary material (stacks.iop.org/JNE/14/036023/mmedia).

References

- Ambard M, Rotter S. Support vector machines for spike pattern classification with a leaky integrate-and-fire neuron. *Frontiers Comput Neurosci.* 2012; 6:78.
- Apkarian AV, Hashmi JA, Baliki MN. Pain and the brain: specificity and plasticity of the brain in clinical chronic pain. *Pain.* 2011; 152:S49–64. [PubMed: 21146929]
- Armstrong-James M, George MJ. Bilateral receptive fields of cells in rat SmI cortex. *Exp Brain Res.* 1988; 70:155–65. [PubMed: 3402562]
- Baliki MN, Petre B, Torbey S, Herrmann KM, Huang L, Schnitzer TJ, Fields HL, Apkarian AV. Corticostriatal functional connectivity predicts transition to chronic back pain. *Nat Neurosci.* 2012; 15:1117–9. [PubMed: 22751038]
- Becerra L, Borsook D. Signal valence in the nucleus accumbens to pain onset and offset. *Eur J Pain.* 2008; 12:866–9. [PubMed: 18226937]
- Boyce-Rustay JM, Honore P, Jarvis MF. Animal models of acute and chronic inflammatory and nociceptive pain. *Methods Mol Biol.* 2010; 617:41–55. [PubMed: 20336412]
- Brown EN, Frank LM, Tang D, Quirk MC, Wilson MA. A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J Neurosci.* 1998; 18:7411–25. [PubMed: 9736661]
- Buesing, L., Macke, JH., Sahani, M. *Advances in Neural Information Processing Systems 25.* New York: Curran Associates; 2013. Spectral learning of linear dynamics from generalised- linear observations with application to neural population data.
- Bushnell MC, Ceko M, Low LA. Cognitive and emotional control of pain and its disruption in chronic pain. *Nat Rev Neurosci.* 2013; 14:502–11. [PubMed: 23719569]
- Chen Z. An overview of Bayesian methods for neural spike train analysis. *Comput Intell Neurosci.* 2013; 2013:251905. [PubMed: 24348527]
- Chen, Z., editor. *Advanced State Space Methods for Neural and Clinical Data.* Cambridge: Cambridge University Press; 2015.
- Chen, Z., Barbieri, R., Brown, EN. State-space modeling of neural spike train and behavioral data. In: Oweiss, K., editor. *Statistical Signal Processing for Neuroscience and Neurotechnology.* Amsterdam: Elsevier; 2010. p. 175-218.
- Chen Z, Kloosterman F, Layton S, Wilson MA. Transductive neural decoding for unsorted neuronal spikes of rat hippocampus. *Proc of IEEE Conf on Engineering in Medicine and Biology.* 2012:1310–3.
- Chen Z, Ohara S, Cao J, Vialatte F, Lenz FA, Cichocki A. Statistical modeling and analysis of laser-evoked potentials of electrocorticogram recordings from awake humans. *Comput Intell Neurosci.* 2007; 2007:10479.
- Coghill RC, Sang CN, Maisog JM, Iadarola MJ. Pain intensity processing within the human brain: a bilateral, distributed mechanism. *J Neurophysiol.* 1999; 82:1934–43. [PubMed: 10515983]
- Copits BA, Pullen MY, Gereau RW IV. Spotlight on pain: optogenetic approaches for interrogating somatosensory circuits. *Pain.* 2016; 157:2424–33. [PubMed: 27340912]
- Craig AD. How do you feel? Interoception: the sense of the physiological condition of the body. *Nat Rev Neurosci.* 2002; 3:655–66. [PubMed: 12154366]
- Craig AD. Pain mechanisms: labeled lines versus convergence in central processing. *Ann Rev Neurosci.* 2003; 26:1–30. [PubMed: 12651967]
- Fraser GW, Chase SM, Whitford A, Schwartz AB. Control of a brain-computer interface without spike sorting. *J Neural Eng.* 2009; 6:055004. [PubMed: 19721186]
- Fuchs PN, Peng YB, Boyette-Davis JA, Uhelski ML. The anterior cingulate cortex and pain processing. *Frontiers Integr Neurosci.* 2014; 8:35.
- Gu L, Uhelski ML, Anand S, Romero-Ortega M, Kim Y, Fuchs PN, Mohanty SK. Pain inhibition by optogenetic activation of specific anterior cingulate cortical neurons. *PLoS One.* 2015; 10:e0117746. [PubMed: 25714399]

- Haupt SS, Spengler F, Husemann R, Dinse HR. Receptive scatter, topography and map variability in different layers of the hindpaw representation of the rat somatosensory cortex. *Exp Brain Res.* 2004; 155:485–99. [PubMed: 14745463]
- Iyer SM, Vesuna S, Ramakrishnan C, Huynh K, Young S, Berndt A, Lee SY, Gorini CJ, Deisseroth K, Delp SL. Optogenetic and chemogenetic strategies for sustained inhibition of pain. *Sci Rep.* 2016; 6:30570. [PubMed: 27484850]
- Jaggi AS, Jain V, Singh N. Animal models of neuropathic pain. *Fund Clin Pharmacol.* 2011; 25:1–28.
- Khan ME, Aravkin AY, Friedlander MP, Seeger M. Fast dual variational inference for non-conjugate latent Gaussian models. *Proc of Int Conf on Machine Learning.* 2013
- Kim KJ, Yoon YW, Chung JM. Comparison of three rodent neuropathic pain models. *Exp Brain Res.* 1997; 113:200–6. [PubMed: 9063706]
- Kloosterman F, Layton S, Chen Z, Wilson MA. Bayesian decoding of unsorted spikes in the rat hippocampus. *J Neurophysiol.* 2014; 111:217–27. [PubMed: 24089403]
- Kuo CC, Yen CT. Comparison of anterior cingulate and primary somatosensory neuronal responses to noxious laser-heat stimuli in conscious, behaving rats. *J Neurophysiol.* 2005; 94:1825–36. [PubMed: 16105955]
- Lee M, Manders TR, Eberle SE, Su C, D'amour J, Yang R, Lin HY, Deisseroth K, Froemke RC, Wang J. Activation of corticostriatal circuitry relieves chronic neuropathic pain. *J Neurosci.* 2015; 35:5247–59. [PubMed: 25834050]
- Lewicki MS. A review of methods for spike sorting: the detection and classification of neural action potentials. *Network.* 1998; 9:R53–78. [PubMed: 10221571]
- Li XY, Wang N, Wang YJ, Zuo ZX, Koga K, Luo F, Zhuo M. Long-term temporal imprecision of information coding in the anterior cingulate cortex of mice with peripheral inflammation or nerve injury. *J Neurosci.* 2014; 34:10675–87. [PubMed: 25100600]
- Macke, JH., Buesing, L., Cunningham, JP., Yu, BM., Shenoy, KV., Sahani, M. *Advances in Neural Information Processing Systems* 24. New York: Curran Associates; 2012. Empirical models of spiking in neural populations.
- Macmillan, NA., Creelman, CD. *Detection Theory: a User's Guide.* New York: Psychology Press; 2004.
- Mogil JS. Animal models of pain: progress and challenges. *Nat Rev Neurosci.* 2009; 10:283–94. [PubMed: 19259101]
- Paninski L, Ahmadian Y, Ferreira DG, Koyama S, Rad KR, Vidne M, Vogelstein JT, Wu W. A new look at state-space models for neural data. *J Comput Neurosci.* 2009; 29:107–26. [PubMed: 19649698]
- Perl ER. Ideas about pain, a historical view. *Nat Rev Neurosci.* 2007; 8:71–80. [PubMed: 17180164]
- Pfau, D., Pnevmatikakis, EA., Paninski, L. *Advances in Neural Information Processing Systems.* New York: Curran Associates; 2014. Robust learning of low-dimensional dynamics from large neural ensembles.
- Prescott SA, Ma Q, Koninck YD. Normal and abnormal coding of somatosensory stimuli causing pain. *Nat Neurosci.* 2014; 17:183–91. [PubMed: 24473266]
- Sewards TV, Sewards MA. The medial pain system: neural representations of the motivational aspect of pain. *Brain Res Bull.* 2002; 59:163–80. [PubMed: 12431746]
- Sewards TV, Zherdin A, Tiemann L, Plant C, Ploner M. Decoding an individual's sensitivity to pain from the multivariate analysis of EEG data. *Cerebral Cortex.* 2012; 22:1118–23. [PubMed: 21765182]
- Shimazaki H, Amari S, Brown EN, Gruen S. State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data. *PLoS Comput Biol.* 2012; 8:e1002385. [PubMed: 22412358]
- Smith A, Brown EN. Estimating a state-space model from point process observations. *Neural Comput.* 2003; 15:965–91. [PubMed: 12803953]
- Smith A, Brown EN. State-space algorithms for estimating spike rate functions. *Comput Intell Neurosci.* 2010; 2010:426539.

- Sorkin LS, Wallace MS. Acute pain mechanisms. *Surg Clin North Am.* 1999; 79:213–29. [PubMed: 10352652]
- Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol.* 2005; 93:1074–89. [PubMed: 15356183]
- Ung H, Brown JE, Johnson KA, Younger J, Hush J, Mackey S. Multivariate classification of structural MRI data detects chronic low back pain. *Cereb Cortex.* 2014; 24:1037–44. [PubMed: 23246778]
- Ventura V. Spike train decoding without spike sorting. *Neural Comput.* 2008; 20:923–63. [PubMed: 18085990]
- Vierck CJ, Whitsel BL, Favorov OV, Brown AW, Tommerdahl M. Role of primary somatosensory cortex in the coding of pains. *Pain.* 2013; 154:334–44. [PubMed: 23245864]
- Vogt BA. Pain and emotion interactions in subregions of the cingulate gyrus. *Nat Rev Neurosci.* 2005; 6:533–44. [PubMed: 15995724]
- Wagner TD, Atlas LY, Lindquist MA, Roy M, Woo CW, Kross E. An fMRI-based neurologic signature of physical pain. *New Engl J Med.* 2013; 368:1388–97. [PubMed: 23574118]
- Wu WY, Liu CY, Tsai ML, Yen CT. Nocifensive behavior-related laser heat-evoked component in the rostral agranular insular cortex revealed using morphine analgesia. *Physiol Behav.* 2016; 154:129–34. [PubMed: 26621057]
- Xu J, Brennan TJ. The pathophysiology of acute pain: animal models. *Curr Opin Anaesthesiol.* 2011; 24:508–14. [PubMed: 21841478]
- Zhang Y, Wang N, Wang JY, Chang JY, Woodward DJ, Luo F. Ensemble encoding of nociceptive stimulus intensity in the rat medial and lateral pain systems. *Mol Pain.* 2011; 7:64. [PubMed: 21864358]

Figure 1. Histology and electrode placement. (A) The electrode tip (in circle) at the end of recording experiments was seen in the rat's ACC. (B) Trace mark (in circle) shows the electrode placed in the rat's S1 hindlimb. Scale bar 1 mm.

Figure 2.

Raster plots and peri-event time histograms (PETHs) of pain-modulated S1 and ACC units (250 mW laser intensity). (A) S1 positive responder. (B) S1 negative responder. (C) ACC positive responder. (D) ACC negative responder. In all plots, spikes are aligned with paw withdrawal. Temporal bin size 50 ms. The waveforms from four channels and total spike counts are shown on the top of rasters. Arrow in the each row of rasters indicates the laser onset of each trial.

Figure 3.

Single-trial decoding analysis in a S1 population (left, 12 units, Dataset rat1-S1-1) and an ACC population (right, 7 units, Dataset rat4-ACC-1). ((A), (D)) Sorted population spike counts. Bin size 50 ms. Color bar indicates spike count, with the dark color representing large spike count. ((B), (E)) Estimated mean Z-score (blue curve) from the univariate latent state. The vertical red lines indicate the paw withdrawal. Horizontal dashed lines mark the thresholds of the significant zone. The shaded area marks the confidence intervals, and the red curve marks the empirical Z-score computed directly from the multi-unit spike count. ((C), (F)) Equivalent p -value derived from the mean Z-score (blue curve) of panels (B) and (E), respectively.

Figure 4.

Single-trial decoding examples in S1 (left: 12 units, Dataset rat3-S1-3) and ACC populations (right: 15 units, Dataset rat5-ACC-2). (A) Sorted population spike counts. Bin size 50 ms. The color bar indicates spike count, with dark color representing large spike count. (B) Estimated mean Z-score (blue curve) from the univariate latent state variable. The vertical red line indicates the paw withdrawal. Horizontal dashed lines mark the thresholds of significant zone. The shaded area marks the confidence intervals, and the red curve marks the empirical Z-score computed directly from the multi-unit spike count.

Figure 5.

Alternative detection strategy for detecting acute pain signals in an ACC population (7 units, Dataset rat4-ACC-1). (A) Sorted population spike counts (same as figure 3(A)). Bin size 50 ms. The color bar indicates spike count, with dark color representing large spike count. (B) Arithmetic mean of absolute Z-scores from all 7 units. (C) Geometric mean of absolute Z-scores from all 7 units. (D) Maximum Z-score derived from all 7 units.

Figure 6.

Demonstration of sequential decoding analysis (7 sorted S1 units, Dataset rat2-S1-2) in a 3 min continuous recording. (A) Sorted population spike counts. Time 0 indicates the onset of recording time. (B) Estimated mean Z-score (blue curve) from the univariate latent state variable. The vertical red lines indicate the paw withdrawal. Shaded area marks the confidence intervals. The '□' symbol denotes the false negative for significance criterion of $p < 0.05$.

Figure 7.

Demonstration of sequential MUA decoding analysis from a S1 population (5 tetrodes, Dataset rat2-S1-2) in a 3 min continuous recording (same period as figure 6). (A) Unsorted MUA spike counts from tetrodes. Time 0 indicates the onset of recording time. (B) Estimated mean Z-score (blue curve) from the univariate latent state variable. The shaded area marks the confidence intervals. The vertical red lines indicate the paw withdrawal. The '□' symbol denotes the false negative and the '*' symbol denotes the false positive for significance criterion of $p < 0.05$. (C) Zoom-in duration of panel (B) between 195 and 225 s.

Figure 8.

Snapshot demonstration of sequential MUA decoding analysis from an ACC population (6 tetrodes, Dataset rat5-ACC-2). (A) Unsorted MUA spike counts from tetrodes. Time 0 indicates the onset of recording time. (B) Estimated mean Z-score (blue curve) from the univariate latent state variable. The shaded area marks the confidence intervals. The vertical red lines indicate the paw withdrawal. The '□' symbol denotes the false negative and the '*' symbol denotes the false positive for significance criterion of $p < 0.05$.

Figure 9.

Sensitivity and specificity analyses. ROC curves and AUROC statistics derived from (A) Dataset rat1-S1-1 and (B) Dataset rat4-ACC-1, both based on sorted units. In the S1 example, the optimal threshold (the point closest to the upper left corner) $p = 0.05$ yields a true positive rate of 95% and a false positive rate 5% (average accuracy 95%). In the ACC example, the optimal threshold $p = 0.05$ yields a true positive rate of 100% and a false positive rate 18.8% (average accuracy 90.6%).

Figure 10.

The scatter plot between the maximum (absolute value) of Z-score and the percentage of modulated units. (A) S1 dataset rat3-S1-3 ($n = 30$ trials). (B) ACC dataset rat6-ACC-3 ($n = 32$ trials). The numbers in each plot show the Spearman's rank correlation R and associated p -value.

Figure 11.

Comparison of detection accuracy (true positive rate) with random subset of neurons being used (Dataset rat1-S1-1). The error bar shows the SD from 50 Monte Carlo runs.

Figure 12. Spontaneous baseline population spike activities. (A) S1 population count within two time-separated 1 min baseline periods. (B) ACC population count within two separated 1 min baseline periods. Note that there is more than a 20 min gap between those two baseline periods.

Figure 13.

Impact of model (non)stationarity on pain signal detection (Dataset rat1-S1-1). ((A)–(D)) Four consecutive (second to fifth) trials. In all figures, the model parameters and noise statistics are estimated from the first single trial of S1 recording. Time 0 marks the onset of withdrawal. The horizontal dashed lines mark the significance threshold of $p = 0.05$. The red (solid/dashed) curve shows the inferred Z-score (mean $\pm$ CI) from an off-line EM algorithm, whereas the magenta curve shows the inferred Z-score from an online forward filtering algorithm. For clarity, the confidence intervals of the online estimate is not shown.

Table 1

Summary of experimental data sets.

Dataset	# tetrodes	# units	# trials analyzed	stimulus (mW)
rat1-S1-1	5	12	40	150
rat2-S1-2	5	7	16	250
rat3-S1-3	6	12	30	250
rat3-S1-4	1	3	25	250
rat3-S1-5	4	10	33	250
rat4-ACC-1	3	7	16	150
rat5-ACC-2	6	15	31	250
rat6-ACC-3	5	11	32	250
rat6-ACC-4	6	8	32	250
rat7-ACC-5	4	6	29	250

Table 2

Predictive likelihood comparison for different dimensionality of latent variable $\mathbf{z}^k \in \mathbb{R}^m$ for $m = 1, \dots, 4$. In this dataset (rat1-S1-1), the first half of trials were used for model identification, and the second half of trials were used for testing. In each row, the best performance is marked in bold font. In each column, the best performance is marked with *.

	PLDS_{variational}	PLDS_{Laplace}
$m = 1$	-41 581	-41 616
$m = 2$	-41 010*	-41 071
$m = 3$	-41 058	-40 814
$m = 4$	-41 063	-40 731*

Table 3

Summary of the AUROC statistics in single-trial decoding.

Dataset	Sorted SUA	Unsorted MUA
rat1-S1-1	0.983	0.920
rat2-S1-2	0.988	0.988
rat3-S1-3	0.958	0.906
rat3-S1-4	0.851	0.728
rat3-S1-5	0.965	0.917
rat4-ACC-1	0.963	0.846
rat5-ACC-2	0.957	0.872
rat6-ACC-3	0.887	0.859
rat6-ACC-4	0.901	0.874
rat7-ACC-5	0.904	0.879

Table 4

Comparison of true positive rates based on sorted single-unit activity (SUA) or unsorted multi-unit activity (MUA). The two ratios in the second and third columns were derived from using significance criterion $p = 0.05$ and $p = 0.001$, respectively. The best result in each row is marked in bold font.

Dataset	Sorted SUA		Unsorted MUA		Stimulus (mW)
	$p = 0.05$	$p = 0.001$	$p = 0.05$	$p = 0.001$	
rat3-S1-3	27/30	25/30	23/30	19/30	250
rat3-S1-4	16/25	11/25	6/25	4/25	250
rat3-S1-5	31/33	26/33	24/33	20/33	250
rat6-ACC-3	23/32	12/32	22/32	13/32	250
rat6-ACC-4	27/32	15/32	24/32	14/32	250
rat7-ACC-5	24/29	18/29	23/29	16/29	250