

The 36th New England Statistics Symposium

STATISTICAL MODELS FOR NEURAL SPIKING DATA

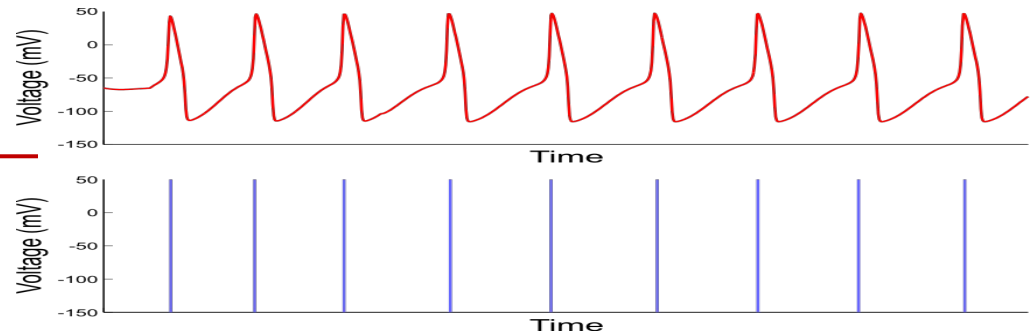
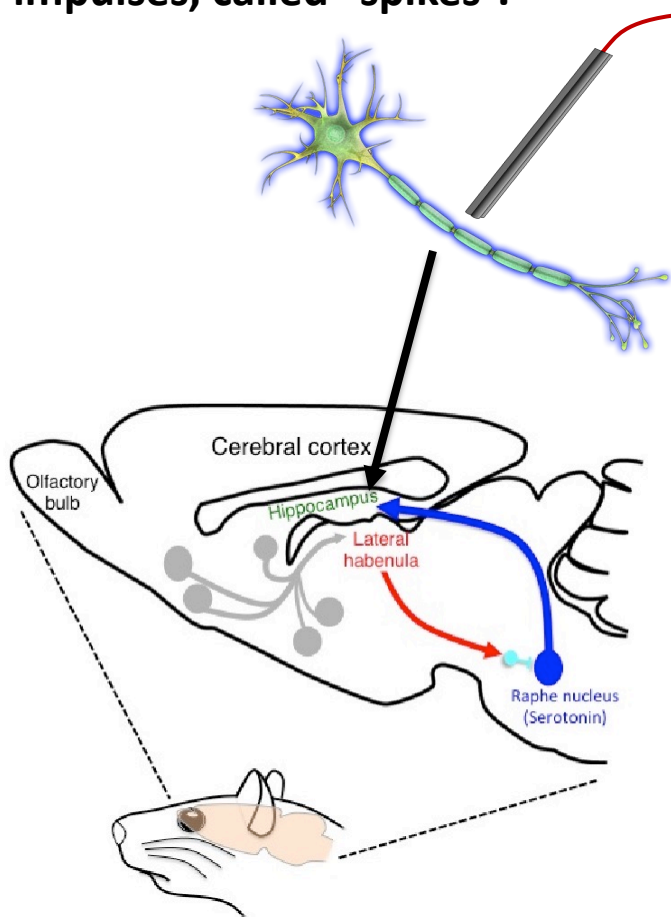
Uri Eden

**Department of Mathematics and Statistics
Boston University**

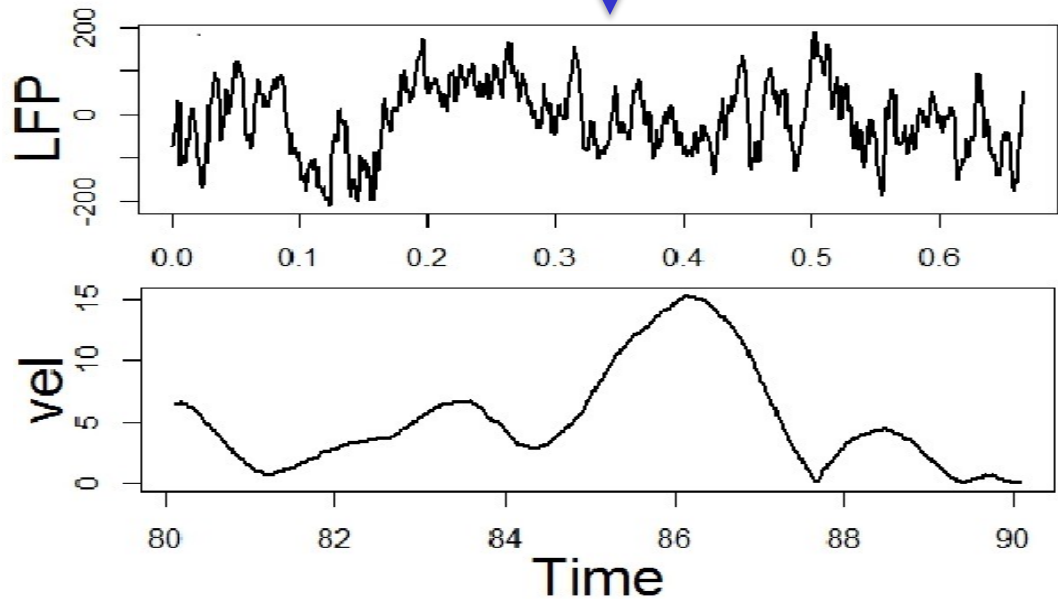
June 3, 2023

Neuroscience experiments require integrating information from multiple sources

Neurons communicate via stereotyped electrical impulses, called “spikes”.



$10^4 \sim 10^5$ neurons



A classic neuroscience experiment

Simple Cell

Features of some classical experiments

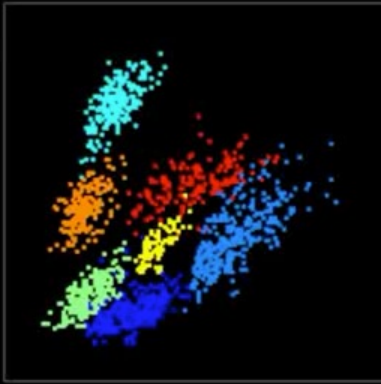
- Recorded activity from an isolated neuron in a single brain area
- Simple stimulus/response paradigm
- All signals are directly observable
- Neural representation clear to the naked eye

A more recent neuroscience experiment

cell activity

behavior

overall



ongoing



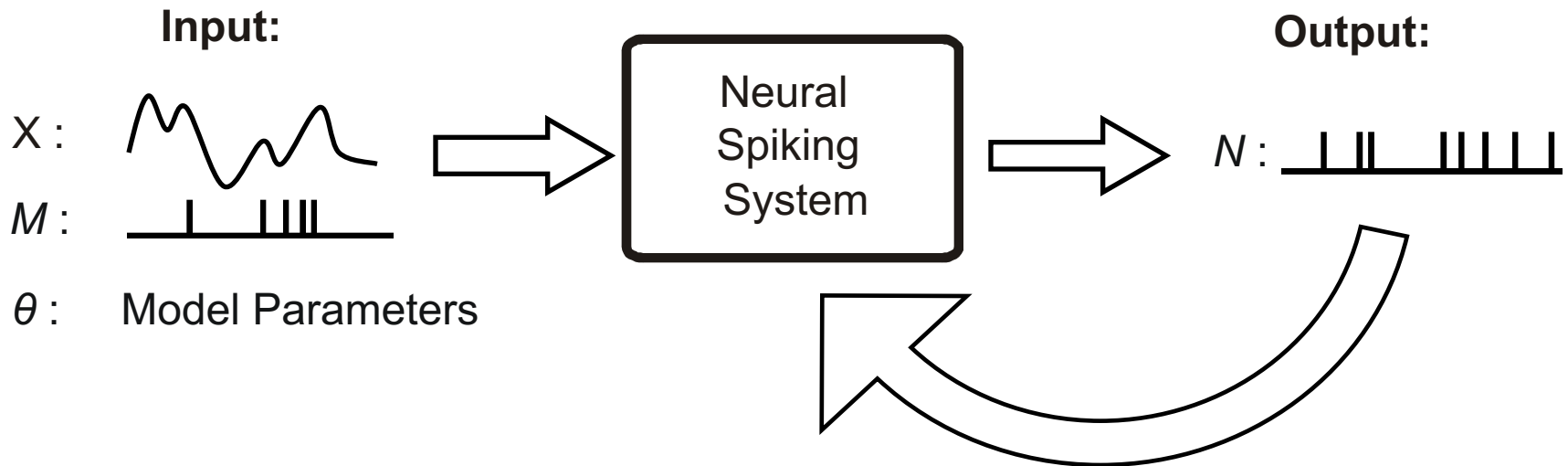
Evolution of experimental and analytic approaches

- Increasing stimulus complexity
 - From simple and experimenter controlled to naturalistic and uncontrolled
 - From single-variate to high dimensional
- Increasing response complexity
 - Single cell responses to coordination of populations within and across multiple brain areas
 - Single response modality to multimodal signals
- Increasing association complexity
 - Stationary models to ones that explicitly capture dynamic representations
 - Directly observable vs hidden dynamics

OUTLINE

- Point process models for individual neurons
- Multivariate point process models for sorted neurons
 - Estimating network structure
- Marked point process modeling
- State space decoding models
 - MATLAB decoding example
- Conclusions

Stochastic Neural Modeling



Point Process System Model: $p(N | x, M, \theta)$

Recursive System Model: $\Pr(\Delta N_t | x_{[0:t]}, M_{[0:t]}, N_{[0:t]}, \theta)$

The Conditional Intensity Function

$$\lambda(t | H_t) = \lim_{\Delta t \rightarrow 0} \frac{\Pr(\text{Spike in } (t, t + \Delta t) | H_t)}{\Delta t}$$

- Unified mathematical construct to model any neural spiking process
- Generalizes Poisson rate function
- Provides building blocks for data likelihood and posterior distributions

$$\Pr(\text{Spike in } (t, t + \Delta t) | H_t) \approx \lambda(t | H_t) \Delta t$$

Point process likelihood

- Likelihood of any spike sequence

$$\rho(s_1, \dots, s_n, N(T) = n) = \prod_{i=1}^n [\lambda(s_i | H_{s_i})] e^{-\int_0^T \lambda(t | H_t) dt}$$

Likelihood of
spike at time s_i



Probability of no
additional spikes in $[0, T]$



Generalized Linear Spiking Models

$$\lambda(t_k|H_k) = \exp \left\{ \theta_0 + \sum_{i=1}^I \alpha_i f_i(\text{Behavioral variables}) \right. \\ \left. + \sum_{j=1}^J \beta_j g_j(\text{Spiking history}) \right. \\ \left. + \sum_{k=1}^K \sum_{c=1}^C \gamma_{k,c} h_{k,c}(\text{Ensemble activity}) \right\}$$

- By selecting appropriate sets of basis functions we can capture arbitrary functional relations.
- Find parameters that maximize likelihood of observed spiking.

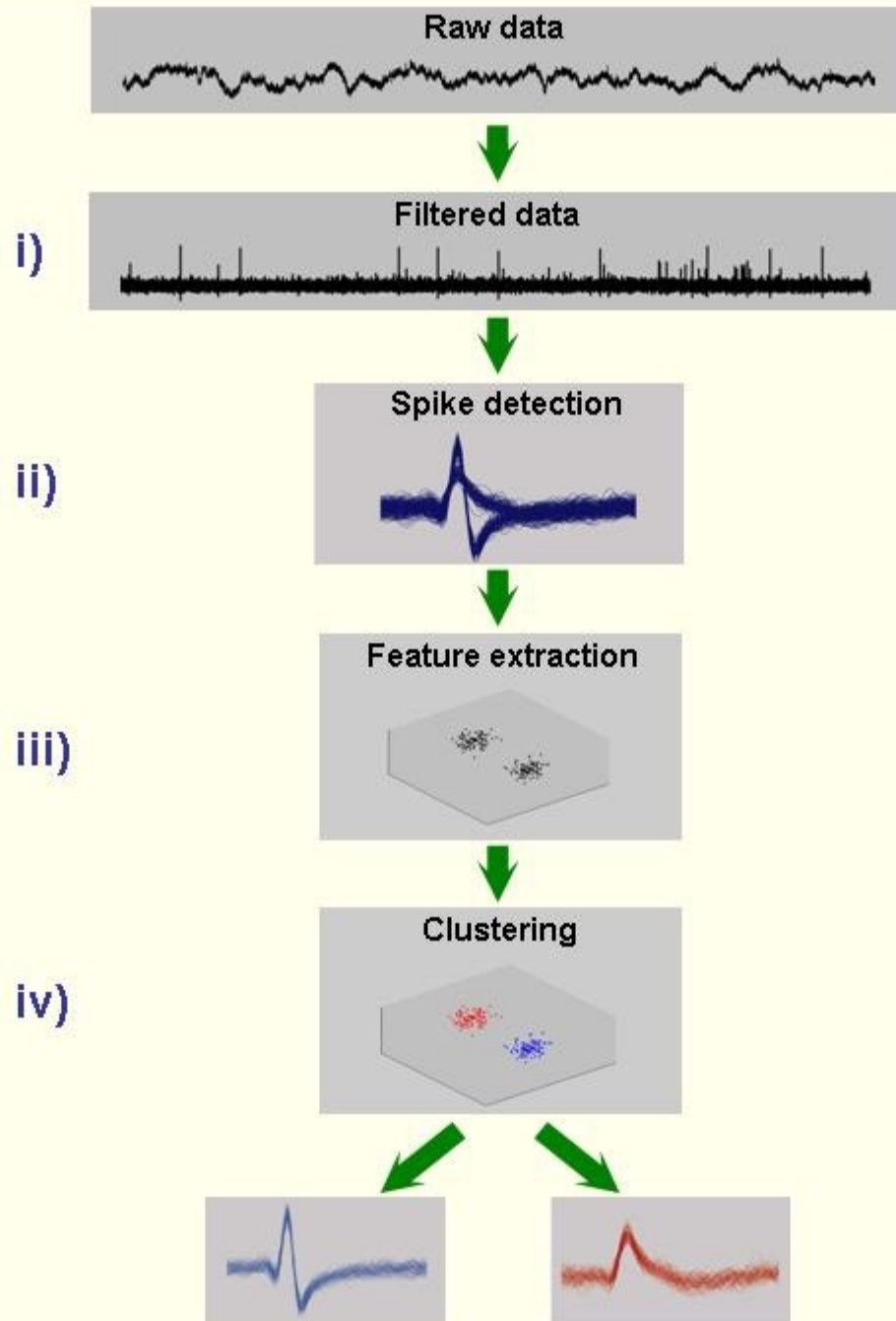
Multivariate point process models for sorted neurons

When

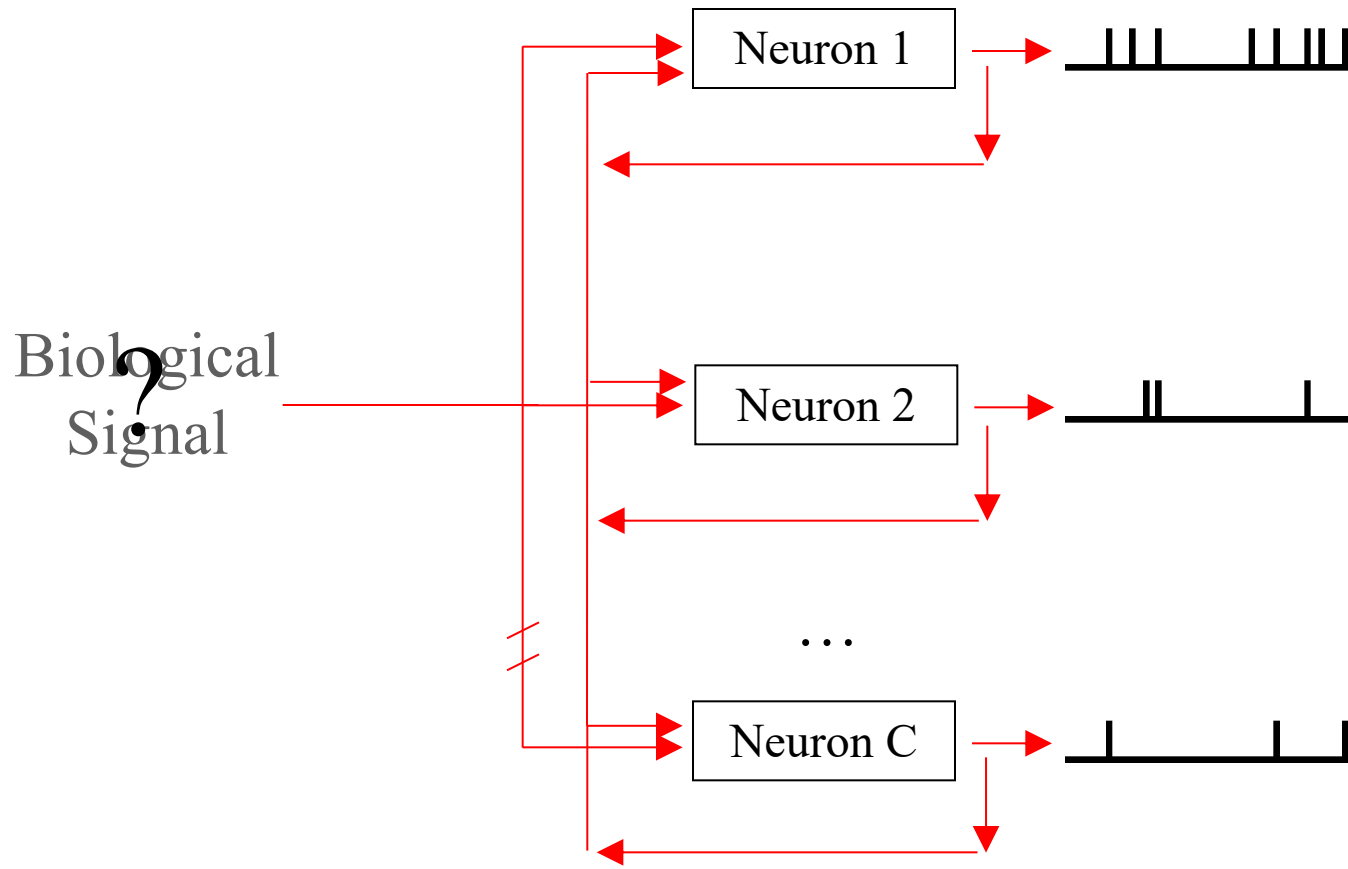
- 1) neural identities of spikes are known with certainty, and
- 2) neurons do not fire exactly simultaneously or with precise temporal spacing,

Then the joint distribution of the population spiking can be modeled by fitting separate models for each neuron as a function of the history of the full ensemble.

Spike Sorting



Multivariate Point Process Model

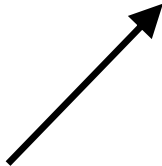


Multivariate point process likelihood

- Likelihood of any spike sequence

$$\rho(s_1^1, \dots, s_{n_1}^1, \dots, s_{n_c}^c, N^1(T) = n_1, \dots, N^c(T) = n_c) \\ = \prod_{c=1}^C \left[\prod_{i=1}^{n_c} [\lambda^c(s_i^c | H_{s_i^c})] e^{-\int_0^T \lambda^c(t | H_t) dt} \right]$$

Likelihood of
spike from neuron
 c at time s_i^c



Probability of no
additional spikes in $[0, T]$



Example: Directed connectivity in crab STG

1. Analysis question: Detect and estimate the structure of coordinated activity between neurons and brain areas.
2. Common approach: Granger causality analysis
 - Linear time-series models for each signal given its own past activity and that of other signals
 - One signal is causal of another in the Granger sense if its inclusion leads to significant improvement in prediction accuracy
 - Note: not the same as physical causality between signals
3. As with many statistical methods, researchers often learn a step-by-step recipe to perform this analysis.



RESEARCH

Open Access

Functional connectivity in a rhythmic inhibitory circuit using Granger causality

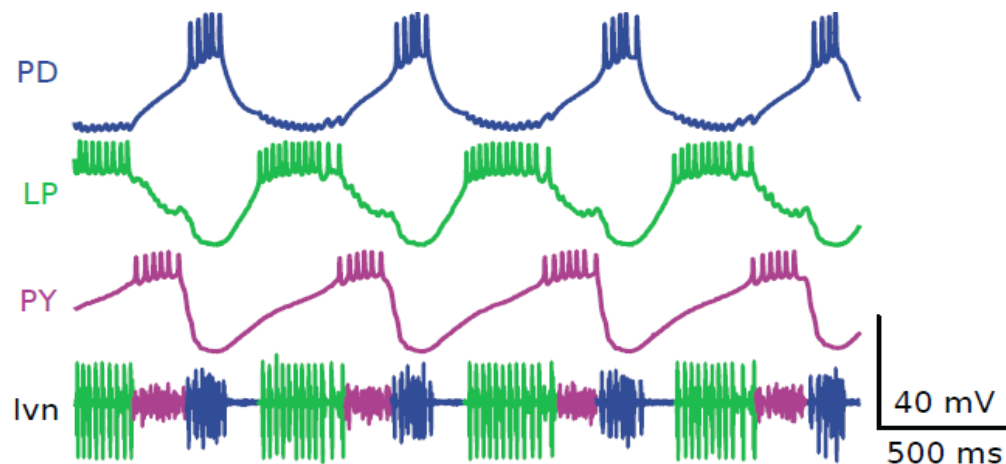
Tilman Kispersky*, Gabrielle J Gutierrez and Eve Marder

Abstract

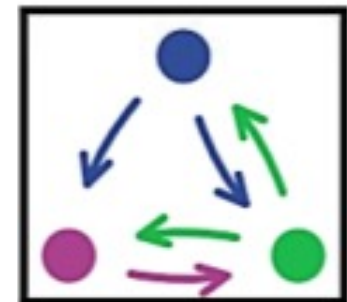
Background: Understanding circuit function would be greatly facilitated by methods that allow the simultaneous estimation of the functional strengths of all of the synapses in the network during ongoing network activity. Towards that end, we used Granger causality analysis on electrical recordings from the pyloric network of the crab *Cancer borealis*, a small rhythmic circuit with known connectivity, and known neuronal intrinsic properties.

Results: Granger causality analysis reported a causal relationship where there is no anatomical correlate because of the strong oscillatory behavior of the pyloric circuit. Additionally, we failed to find a direct relationship between synaptic strength and Granger causality in a set of pyloric circuit models.

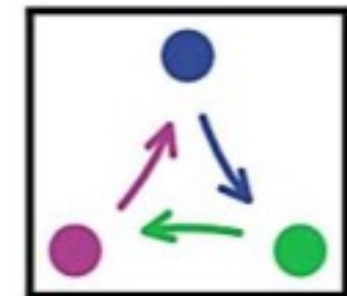
Conclusions: We conclude that the lack of a relationship between synaptic strength and functional connectivity occurs because Granger causality essentially collapses the direct contribution of the synapse with the intrinsic properties of the postsynaptic neuron. We suggest that the richness of the dynamical properties of most biological neurons complicates the simple interpretation of the results of functional connectivity analyses using Granger causality.



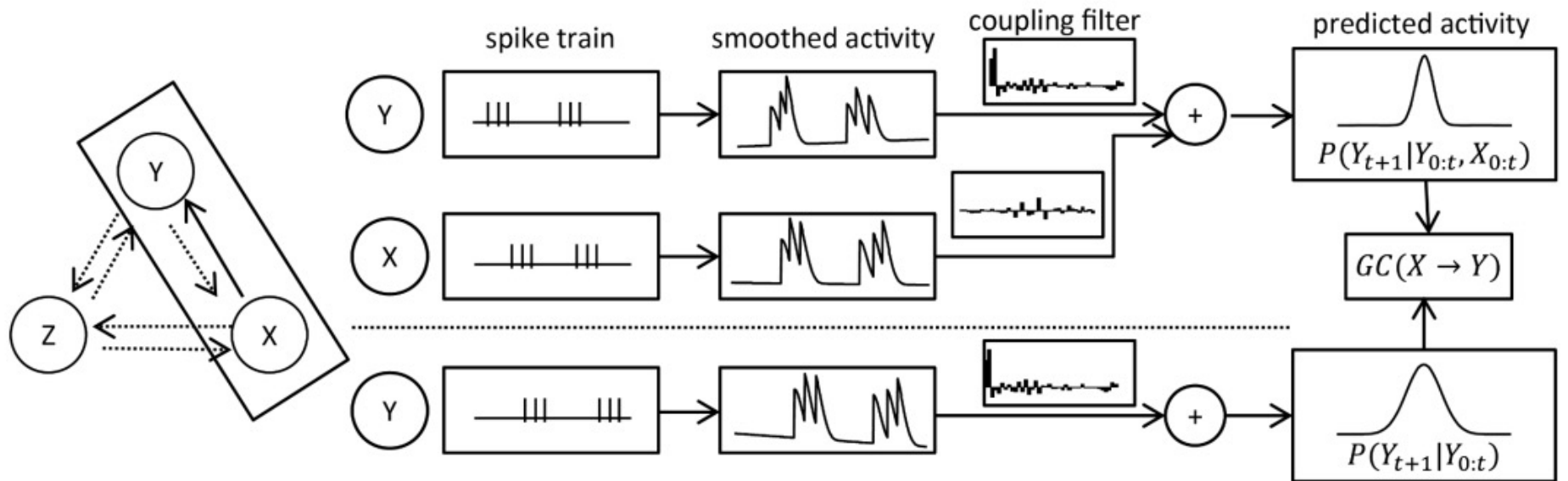
anatomy



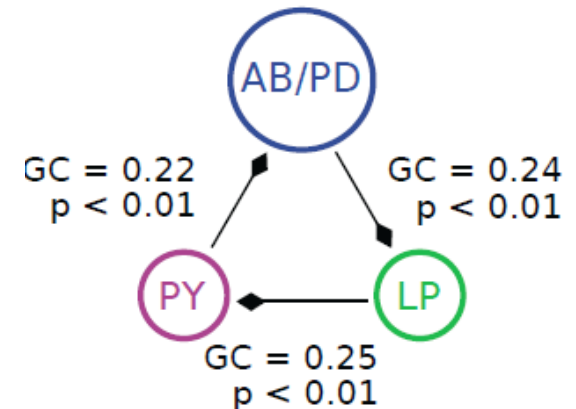
linear Granger



Granger causality (GC) analysis

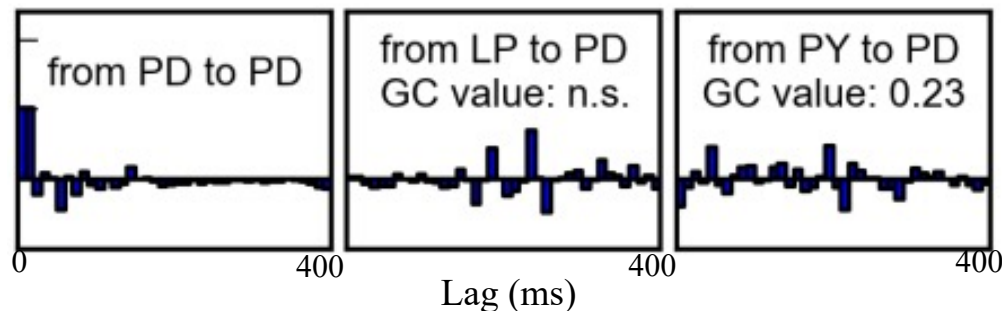


$$Y_t = \sum_{i=1}^I \alpha_i Y_{t-i} + \sum_{j=1}^J \beta_j X_{t-j} + \varepsilon_t$$

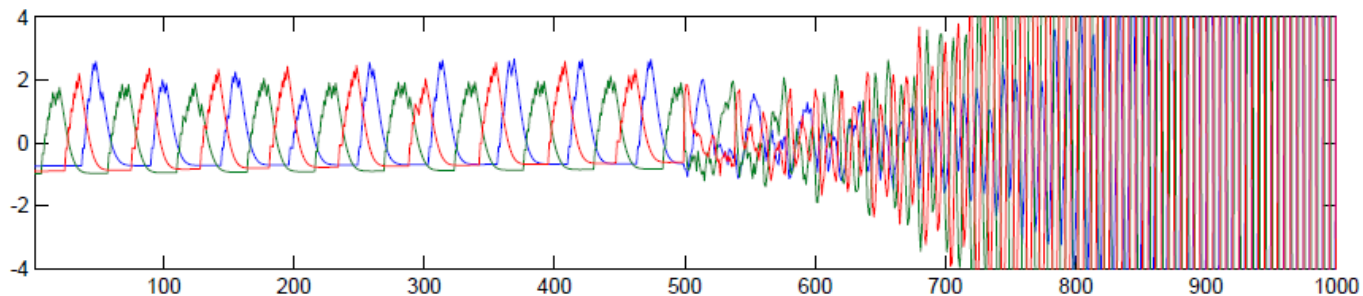


Model interpretation and assessment

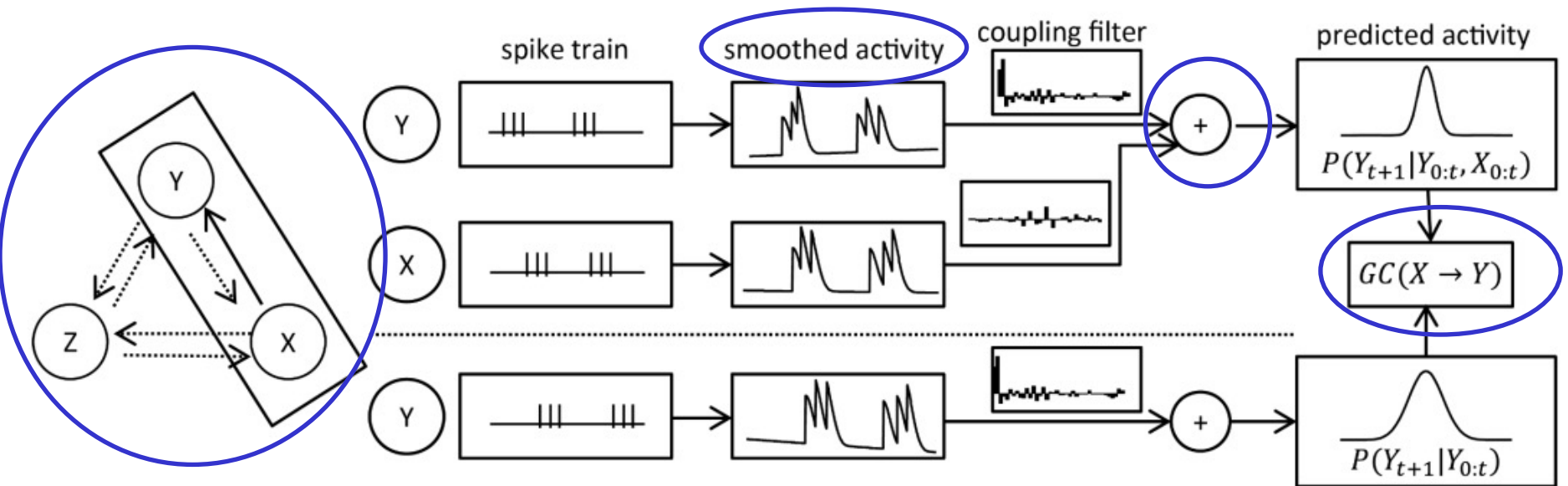
- A statistical model is not a black box. Opportunities to understand and evaluate model components are often missed.
- Example 1: GC coupling parameters don't reflect synaptic influences.



- Example 2: GC model doesn't reproduce rhythm on it's own.

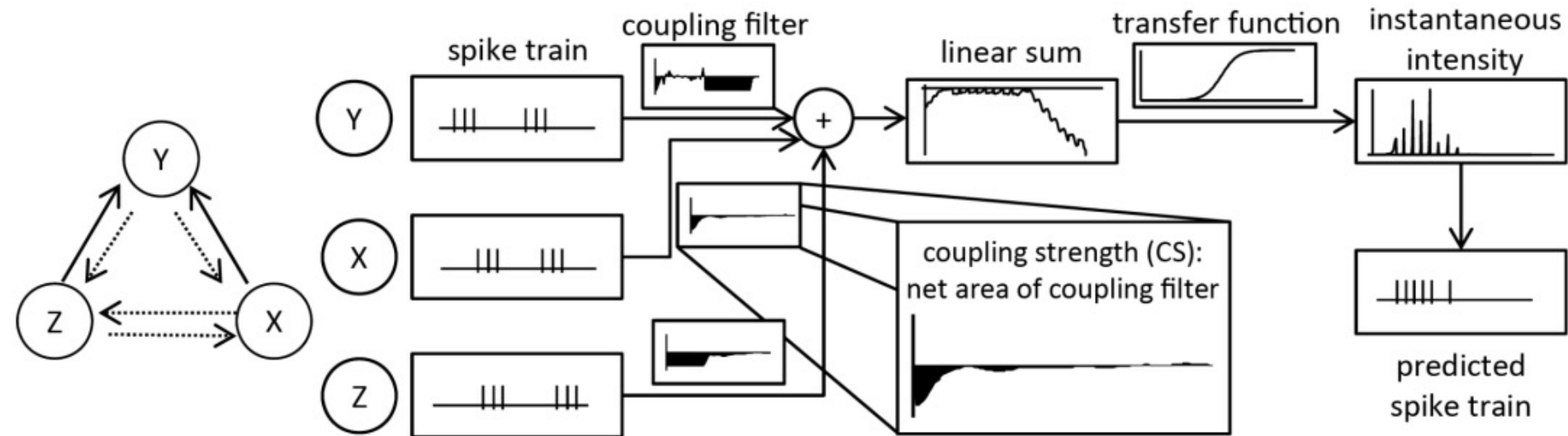


Model refinement



- 1) Use point process model to describe spike trains
- 2) Use nonlinear and multivariate model
- 3) Use physiological significance along with statistical significance

Point process analysis

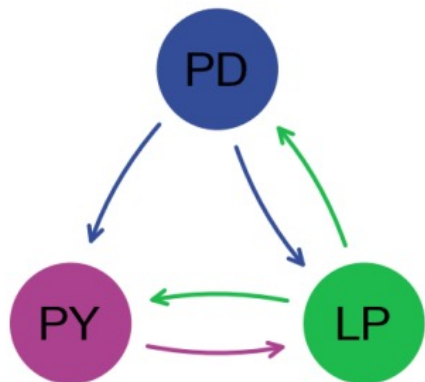


$$\Pr(\text{Y spikes at } t) \propto \exp \left(\sum_{i=1}^I \alpha_i g_i(H_t^Y) + \sum_{j=1}^J \beta_j g_j(H_t^X) + \sum_{k=1}^K \gamma_k g_k(H_t^Z) \right)$$

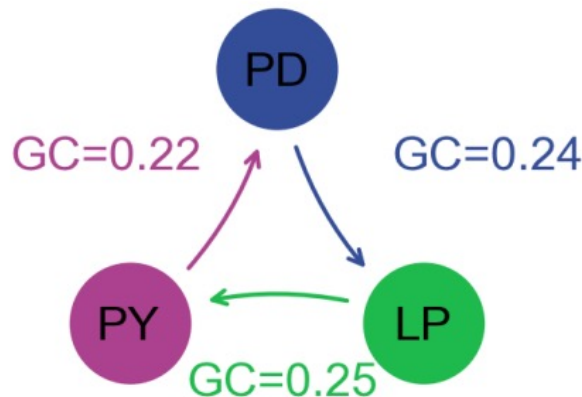
A case study on point process modeling

Estimating functional connectivity in crab STG

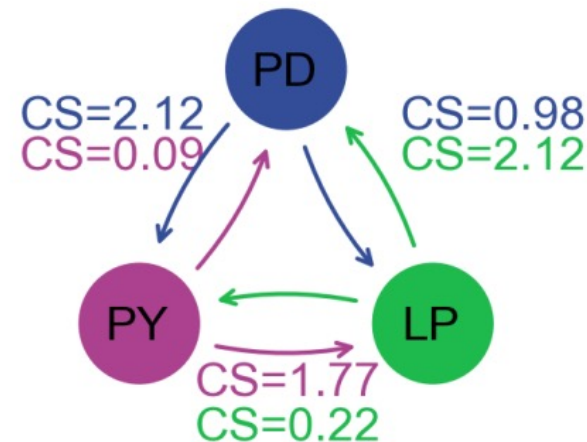
physiological connectivity



Granger causality



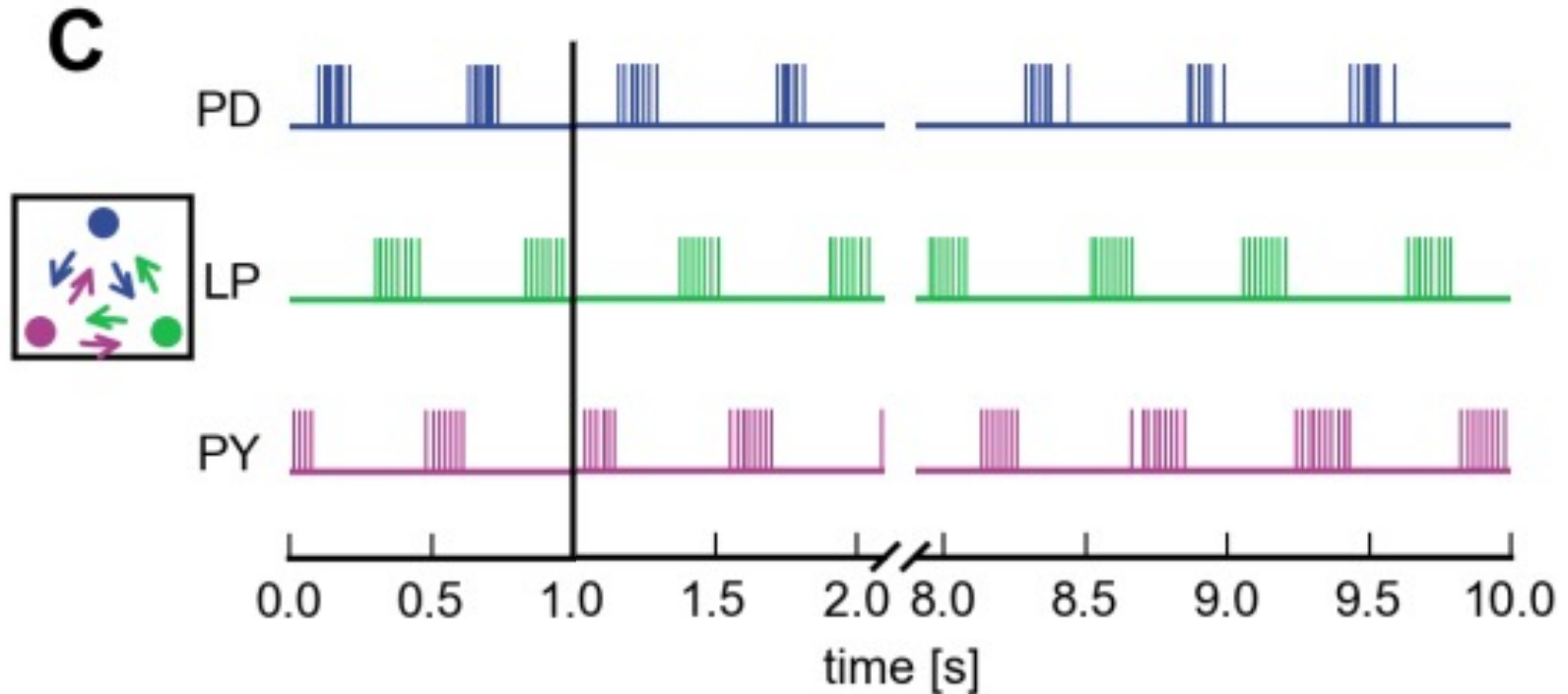
point process model



$$x_t = \sum_{i=1}^I \alpha_i x_{t-i} + \sum_{j=1}^J \beta_j y_{t-j} + \sum_{k=1}^K \gamma_k z_{t-k} + \varepsilon_t$$

$$\Pr(x \text{ spikes at } t) \propto \exp \left(\sum_{i=1}^I \alpha_i g_i(H_t^x) + \sum_{j=1}^J \beta_j g_j(H_t^y) + \sum_{k=1}^K \gamma_k g_k(H_t^z) \right)$$

Point process model reproduces rhythm



$$\Pr(x \text{ spikes at } t) \propto \exp \left(\sum_{i=1}^I \alpha_i g_i(H_t^x) + \sum_{j=1}^J \beta_j g_j(H_t^y) + \sum_{k=1}^K \gamma_k g_k(H_t^z) \right)$$

(Gerhard et al., 2013)

Lessons from this case study

- Importance of using methods that reflect statistical structure of data.
- “All models are wrong, but some are useful” – George E. P. Box
- Powerful goodness-of-fit methods exist to characterize model quality – but are used infrequently.
- Inference methods should be based on interpretable statistics and model parameters.

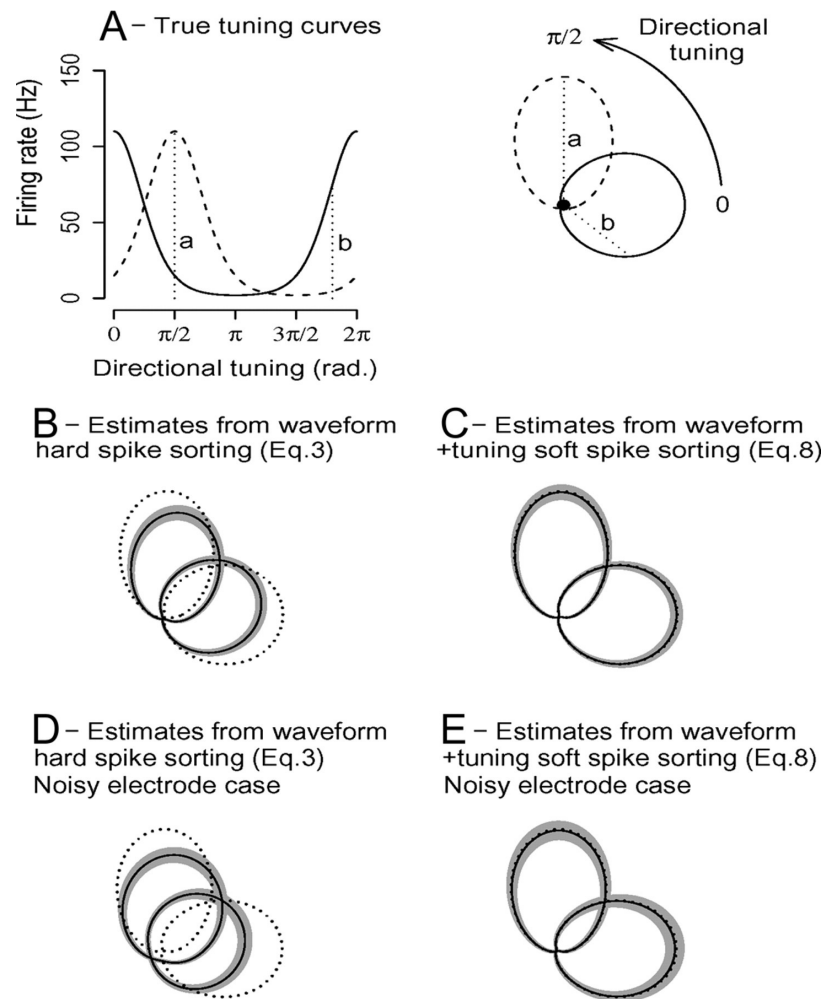
Spike sorting issues

- In many experiments, spikes sorting cannot be performed with complete confidence
 - Uncertainty in sorting is rarely explicitly expressed
 - Propagation of sorting uncertainty to coding models is rarely explored
 - Sorting uncertainty leads to both bias and unmodeled variability in models
 - Data that cannot be easily clustered is often thrown away
- Computational challenge of spike sorting precludes real-time methods (e.g. closed-loop experiments)

Traditional waveform based spike sorting yields biased rate code estimates

Valérie Ventura¹

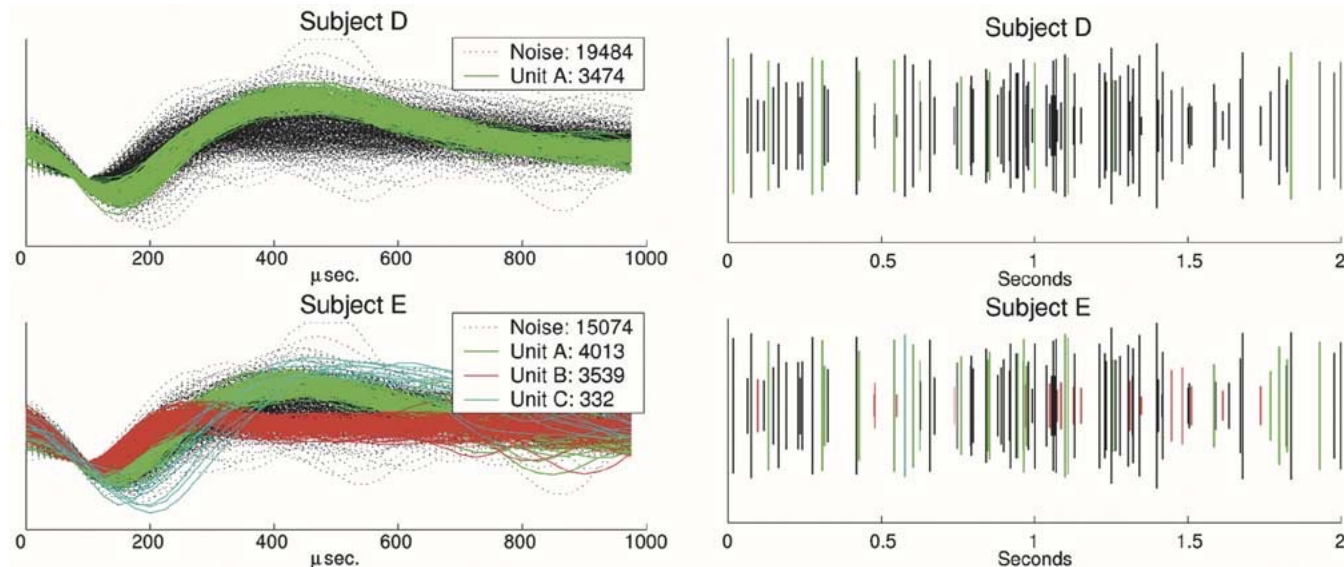
- The standard two-step process of first sorting and then fitting tuning curves leads to biased estimates.
- This, despite the primary motivation cited for a separate sorting step being to prevent a form of bias.
- One alternate approach: soft sorting
- Modeling goals?



On the Variability of Manual Spike Sorting

Frank Wood*, Michael J. Black, *Member, IEEE*, Carlos Vargas-Irwin, Matthew Fellows, and John P. Donoghue, *Member, IEEE*

- Comparisons between trained spike sorters from same lab showed extreme variability in the number of units identified, spikes sorted, and identity assignments.
- In some cases, sorting disregards >80% of threshold crossings



An alternative approach

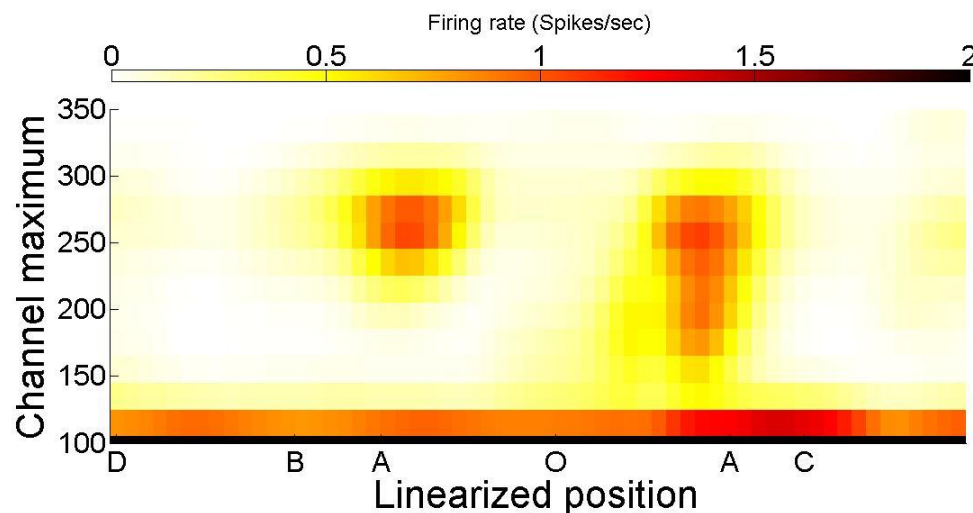
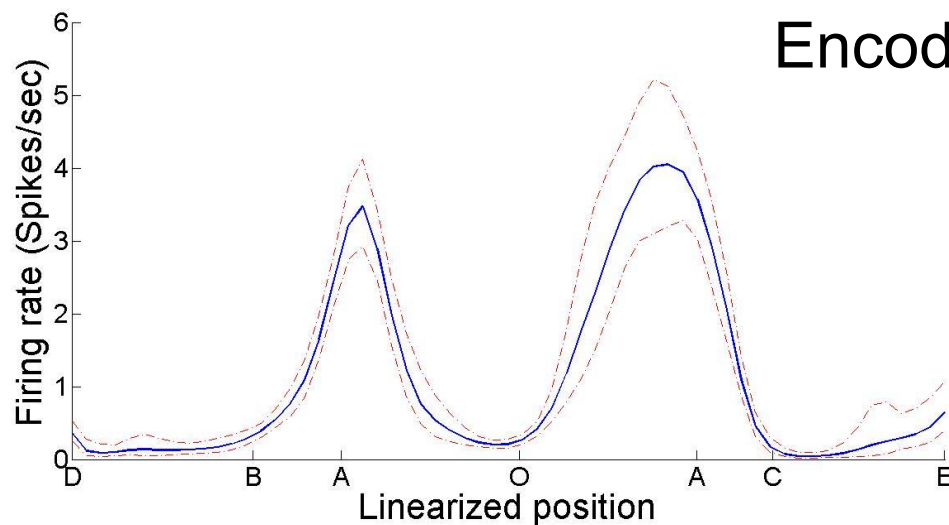
- Model the coding properties of the full neural population directly.
- Encoding model describes the likelihood of observing a spike with a particular waveform as a function of represented signals.
- Let \mathcal{M} be a set of distinguishing waveform features
- A marked point process model is defined by a joint mark-intensity model $\lambda(t, \vec{m})$, for all $\vec{m} \in \mathcal{M}$ defined so that

$$\lim_{\Delta t \rightarrow 0} \frac{\Pr(\text{spike with mark } \vec{m} \in M \text{ in } (t, t + \Delta t))}{\Delta t} = \int_M \lambda(t, \vec{m}) d\vec{m}$$

Marked point process models

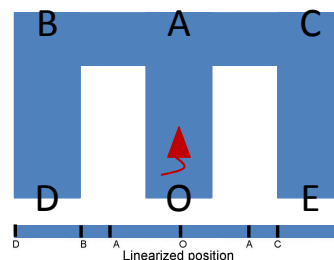
Encoding model with sorted spikes:

$$\hat{\Lambda}^c(x(t))$$



Marked point process encoding model:

$$\hat{\lambda}(x(t), m)$$

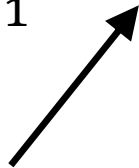


Marked point process models

- Likelihood of any spike sequence

$$\rho((s_1, \vec{m}_1), \dots, (s_n, \vec{m}_n), N(T) = n) \\ = \prod_{i=1}^n [\lambda(s_i, \vec{m}_i | H_{s_i})] e^{-\int_0^T \int_{\mathcal{M}} \lambda(t, \vec{m} | H_t) d\vec{m} dt}$$

Likelihood of
spike at time s_i
with mark \vec{m}_i

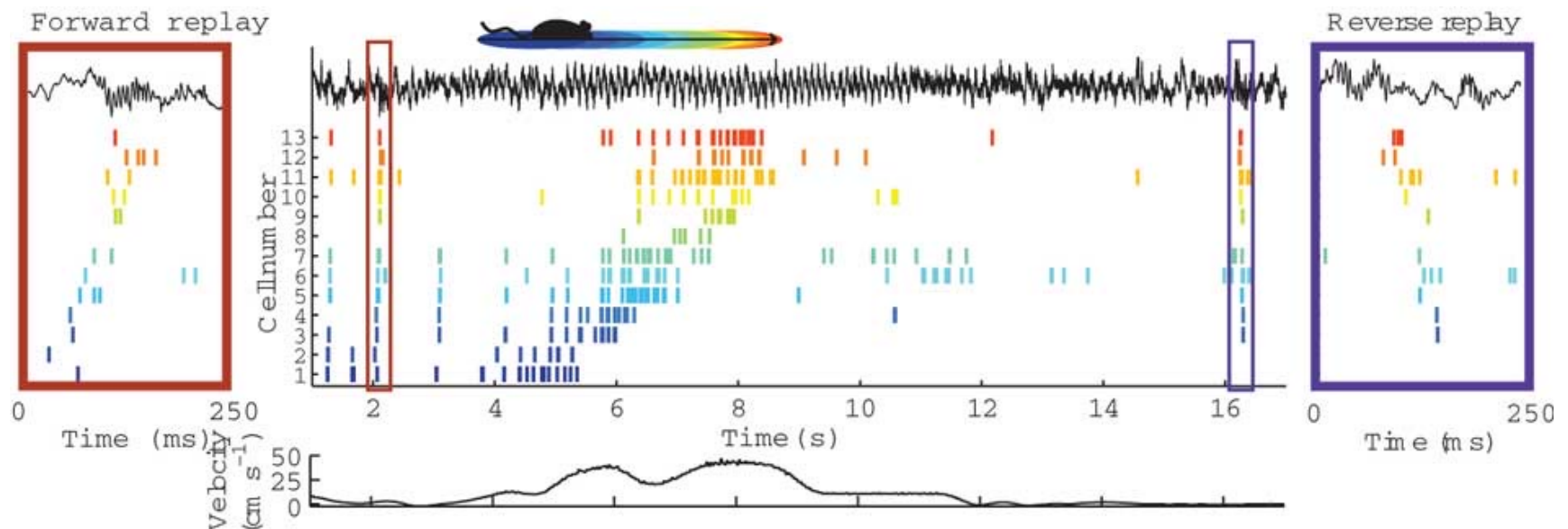


Probability of no
additional spikes in $[0, T]$



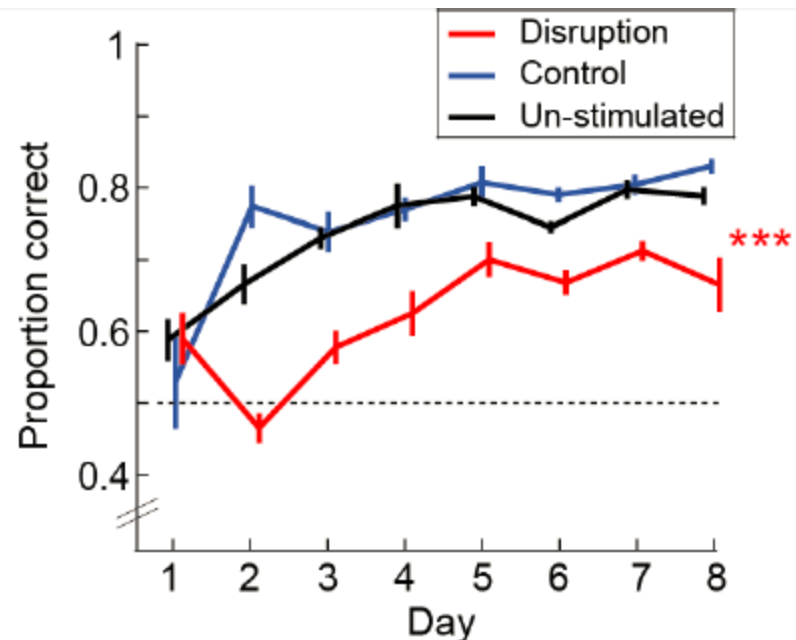
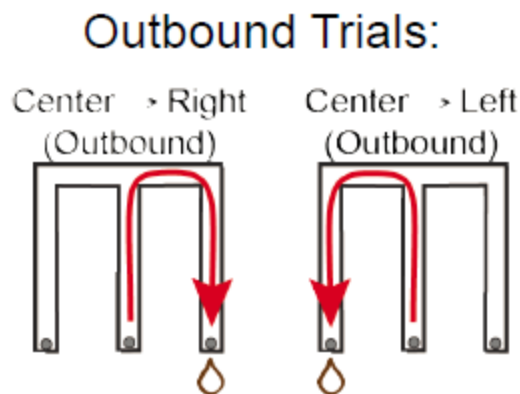
A motivating problem: Awake Hippocampal Replay

- Reoccurrence of sequences of neural activation that also occurred during activity, at a faster time scale
- Important for memory consolidation and learning of navigation tasks



Experimental Challenge

- Interrupting all replay events leads to substantial deficits in navigational learning tasks.
- Ultimate goal: Decode replay content in real time to selectively interrupt events.
- Preliminary goal: Decode movement in real time without spike sorting



Marked point process models

$$\lim_{\Delta t \rightarrow 0} \Delta t^{-1} \Pr[\text{spike with mark in } \mathcal{M} \text{ in } [t, t + \Delta t)] = \int_{\mathcal{M}} \lambda(t, m) dm$$

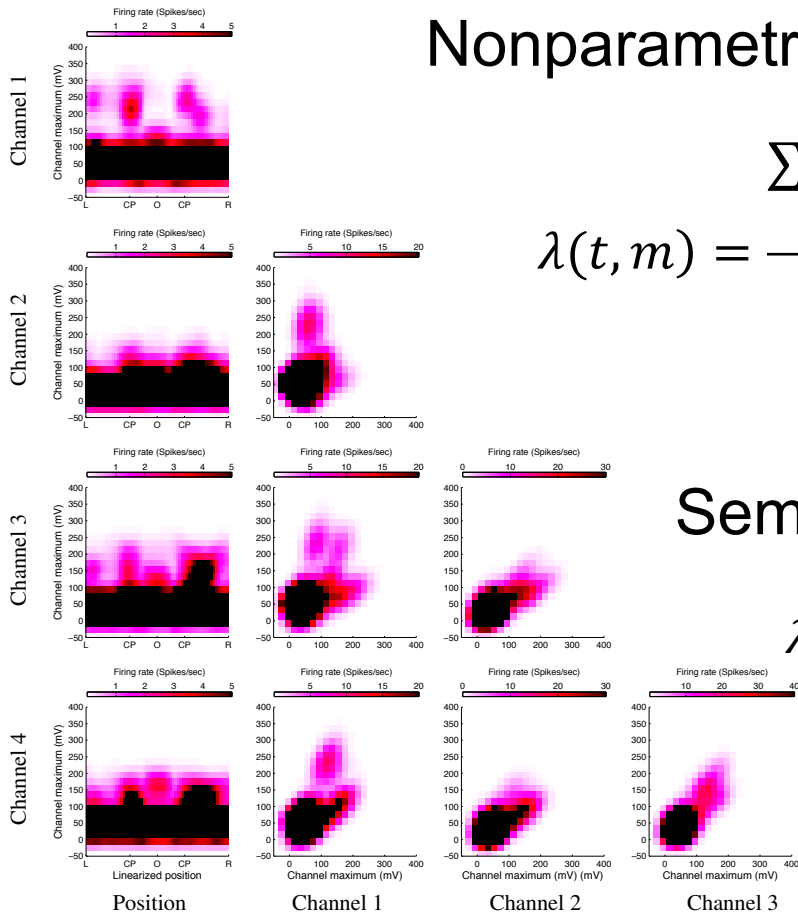
Nonparametric:

$$\lambda(t, m) = \frac{\sum_{i=1}^N \mathcal{K} \left(\frac{x(t) - x_{u_i}}{B_x}, B_m^{-1}(m - m_{u_i}) \right)}{\sum_{j=1}^T K \left(\frac{x(t) - x_j}{b_x} \right)}$$

Semiparametric:

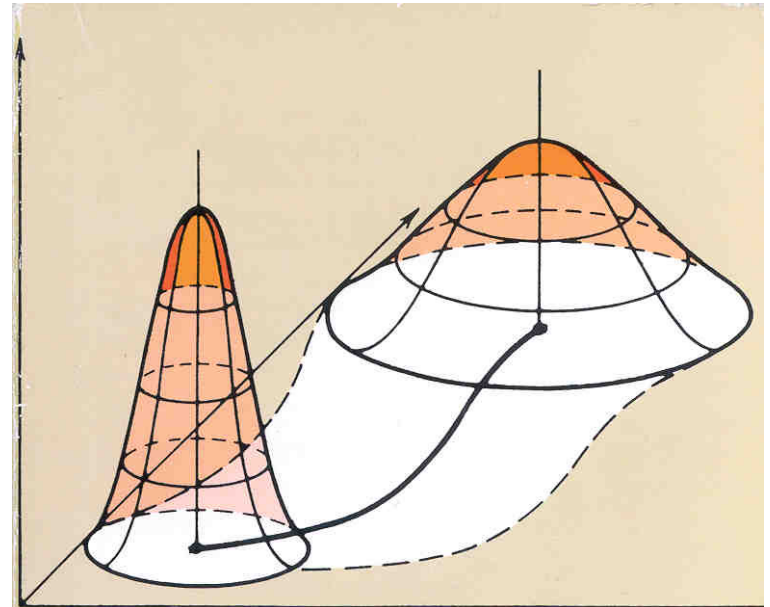
$$\lambda(t, m) =$$

$$\sum_{c=1}^C w_c \mathcal{K} \left(\frac{x(t) - \mu_{x,c}}{\Sigma_{x,c}}, \Sigma_m^{-1}(m - \mu_{m,c}) \right)$$



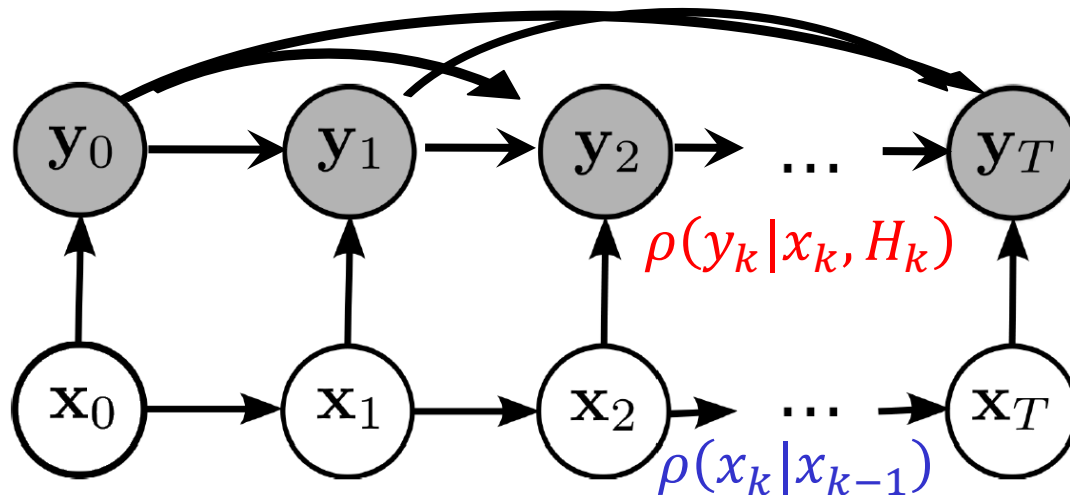
State Estimation

- Neural intensity model characterizes $p(\text{spikes}|\text{state})$
- For state estimation, we must track $p(\text{state}|\text{spikes})$.
 - Get estimators
 - Higher order statistics
- Application of Bayes' rule



The state-space paradigm

- An unobserved state process (x_t) undergoes stochastic dynamics that influence a set of observation processes (y_t)



- Classically, the states represent directly observable features of a behavioral task
 - E.g. Decoding location from ensemble spiking in rat hippocampus
- Recently, states have been used to represent unobserved, low-dimensional dynamics of neural representations

Decoding filter

State model: $\rho(x_k|x_{k-1})$

Observation model:

$$\rho\left(\Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k} | x_k, H_k\right) = \prod_{i=1}^{\Delta N_k} [\lambda(t_k, m_{k_i})] e^{-\Delta t \int_{\mathcal{M}} \lambda(t_k, m) dm}$$

Bayes' rule:

$$\rho\left(x_k | \Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k}, H_k\right) \propto \rho\left(\Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k} | x_k, H_k\right) \rho(x_k | H_k)$$

Chapman-Kolmogorov equation:

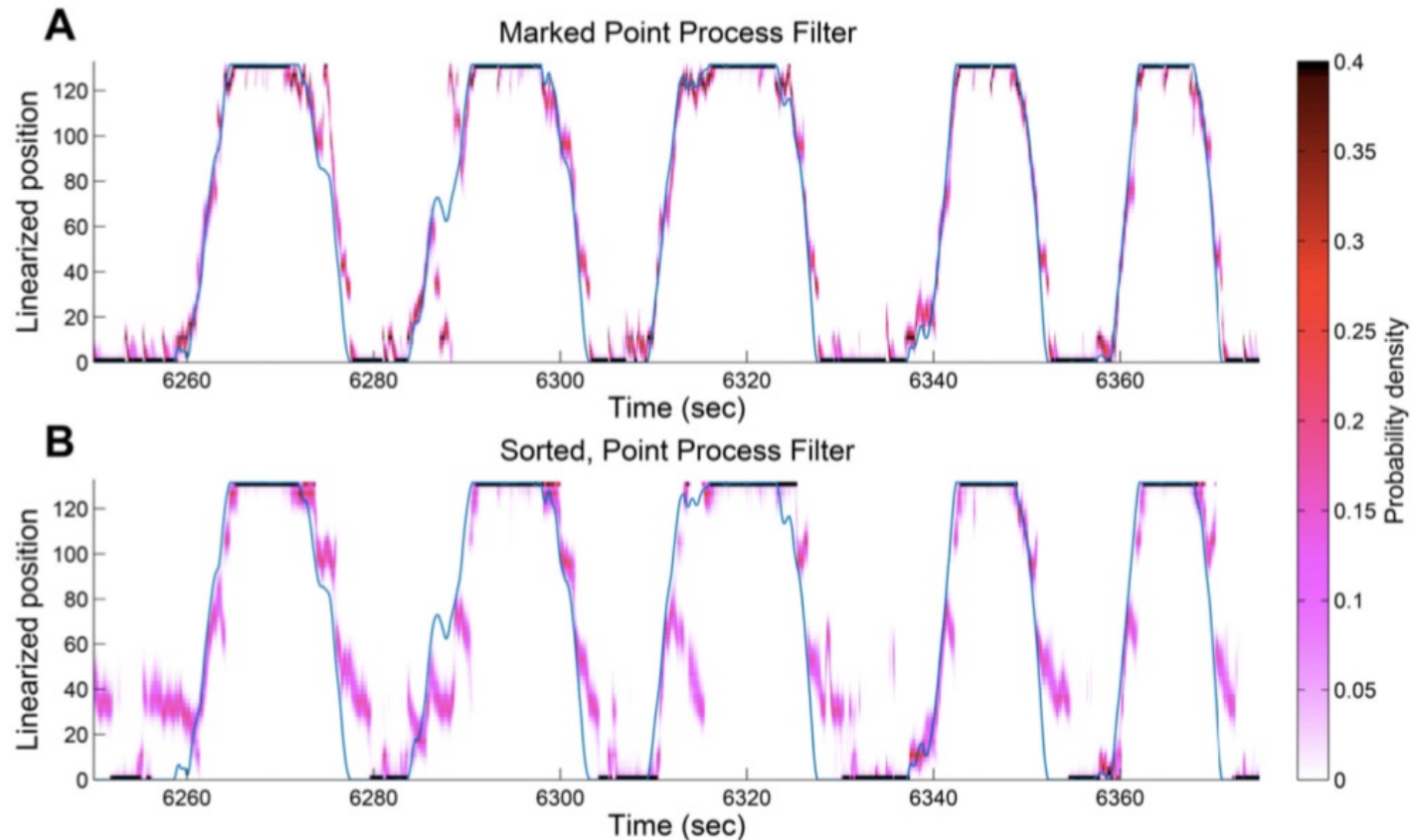
$$\rho(x_k | H_k) = \int \rho(x_k | x_{k-1}) \rho\left(x_{k-1} | \Delta N_{k-1}, \{m_{k-1_i}\}_{i=1}^{\Delta N_{k-1}}, H_{k-1}\right) dx_{k-1}$$

Iterative equation for posterior density:

$$\rho\left(x_k | \Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k}, H_k\right) \propto \rho\left(\Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k} | x_k, H_k\right) \int \rho(x_k | x_{k-1}) \rho\left(x_{k-1} | \Delta N_{k-1}, \{m_{k-1_i}\}_{i=1}^{\Delta N_{k-1}}, H_{k-1}\right) dx_{k-1}$$

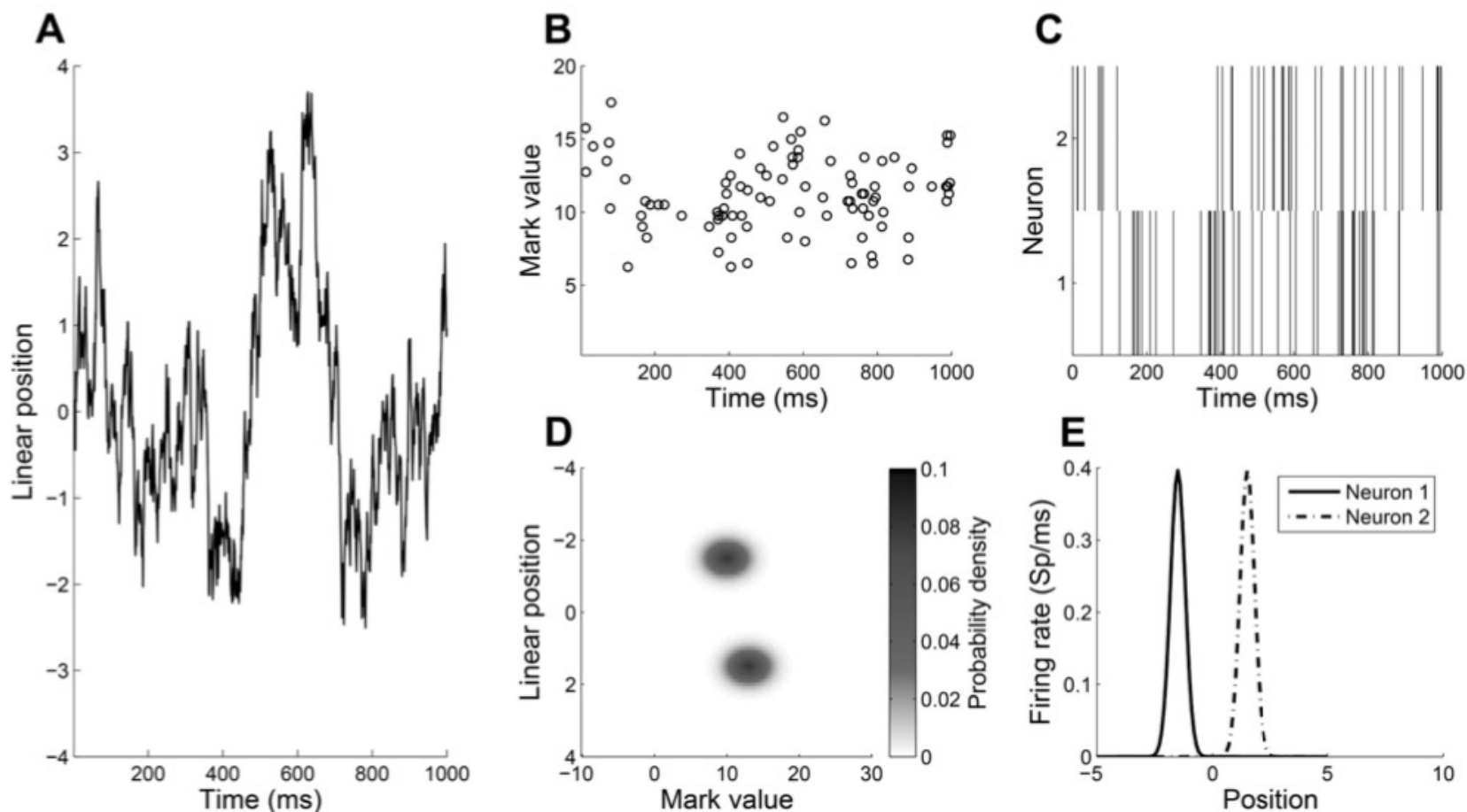
Decoding – model comparison

- Surprisingly, clusterless decoding trajectories were more certain and more accurate



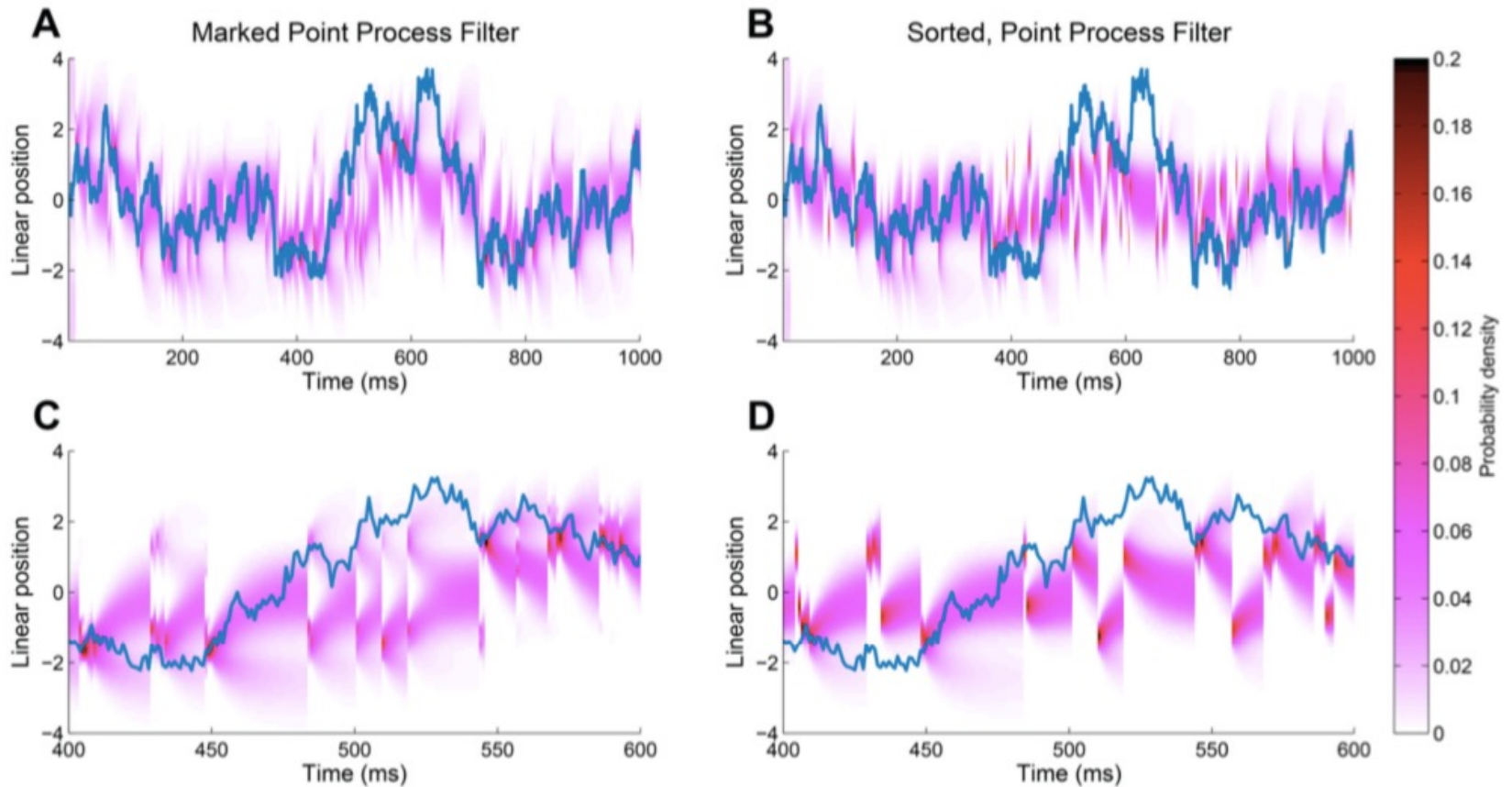
Improvement in decoding

- Simple simulation involving two neurons with overlapping waveform features



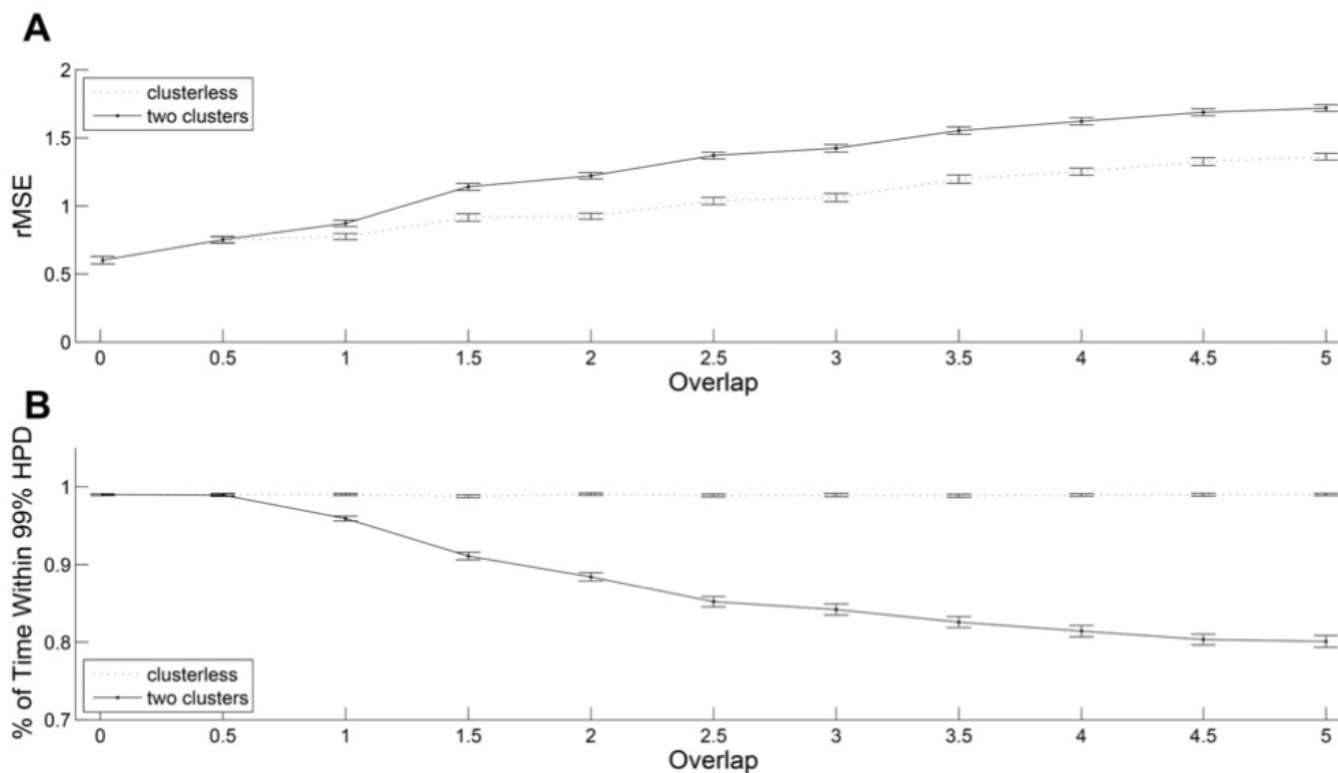
Improvement in decoding

- Simple simulation involving two neurons with overlapping waveform features



Improvement in decoding

- Simple simulation involving two neurons with overlapping waveform features



Clusterless decoding conclusions

- Marked point processes produce real-time decoding algorithms suitable for closed-loop experiments
- Clusterless models eliminate bias imposed by two step (sorting, model fitting) process
- Clusterless models allow us to retain unclassifiable spikes that provide coding information
- Clusterless decoding provides a natural measure for assessing the quality of population-level coding models.

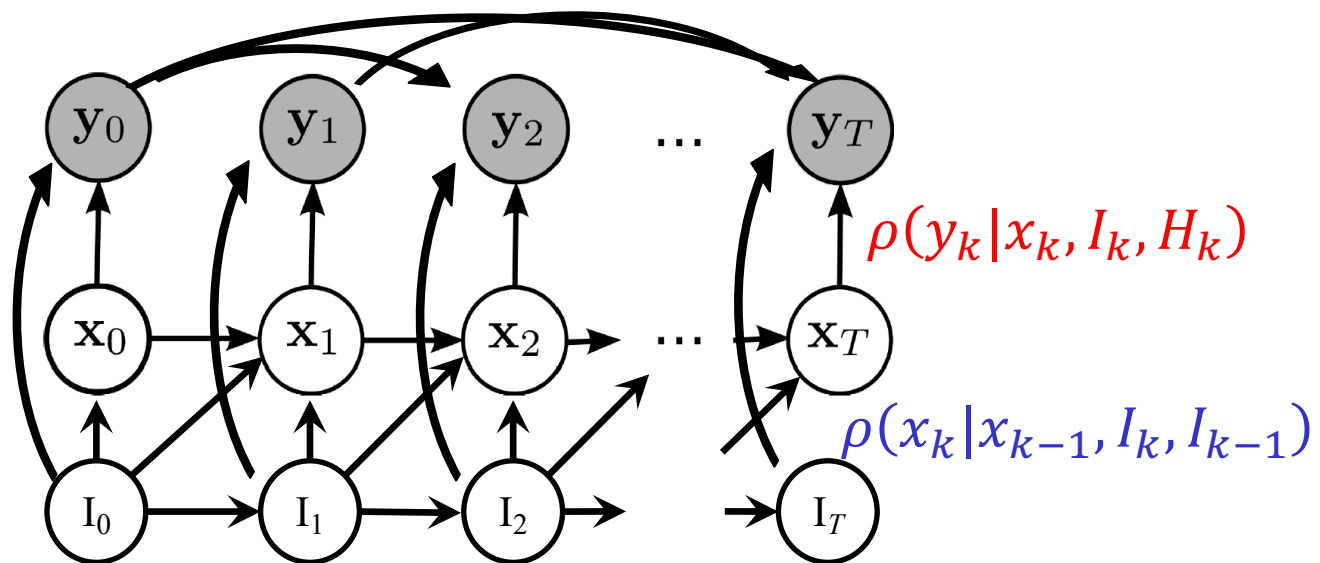
Back to motivating example (if there's time)

Challenge:

Replay relates to a process of reactivation that is not directly observable.

Approach – State-space estimation:

Discrete and continuous latent processes that drive observations are estimated using a filter/smoothing + EM algorithm.



Decoding discrete replay features

Initial condition: $\rho(I, x_0) = \rho(x_0|I)\rho(I)$

State model: $\rho(x_k|x_{k-1}, I)$

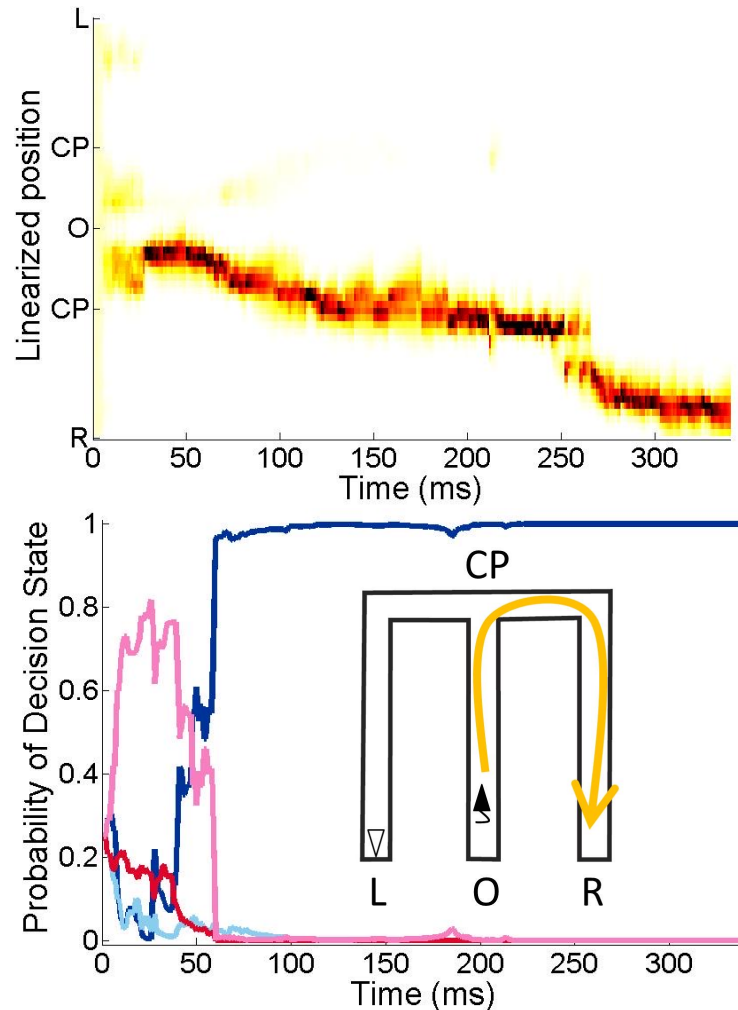
Observation model: $\rho(\Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k} | I, x_k, H_k)$

Discrete state, marked point process filter:

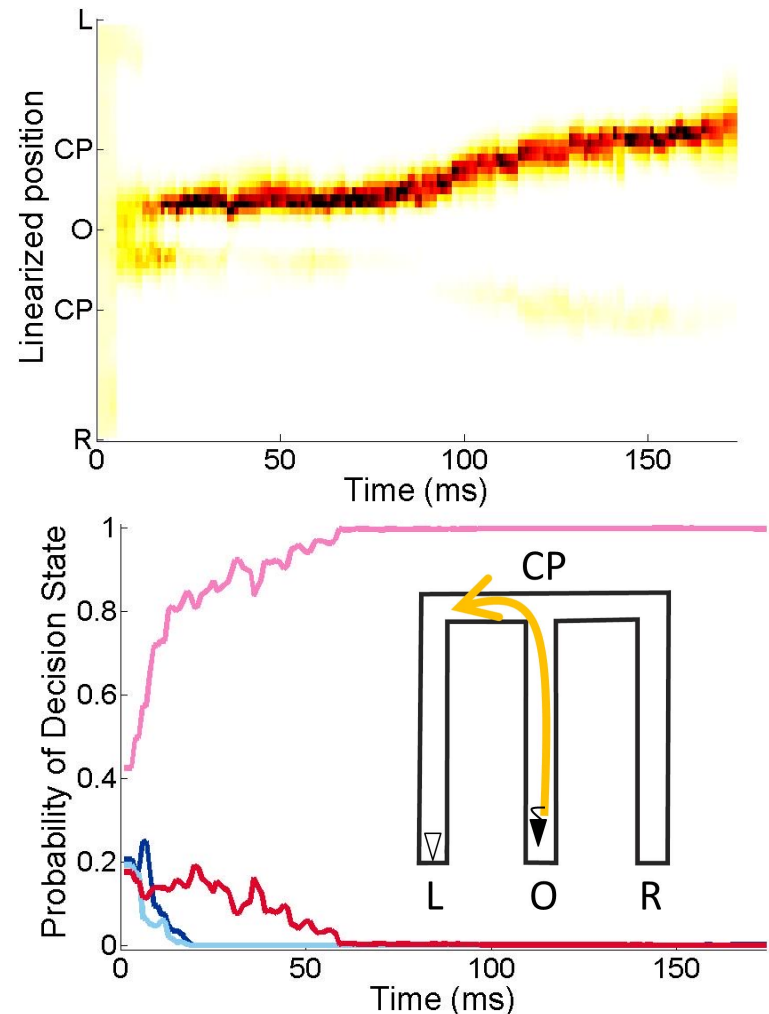
$$\Pr(I|H_{k+1}) \propto \int \rho(I, x_k | \Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k}, H_k) dx_k \propto$$
$$\int \rho(\Delta N_k, \{m_{k_i}\}_{i=1}^{\Delta N_k} | I, x_k, H_k) \int \rho(x_k | x_{k-1}, I) \rho(I, x_{k-1} | H_k) dx_{k-1} dx_k$$

Real time decoding of discrete replay features

Replay 1

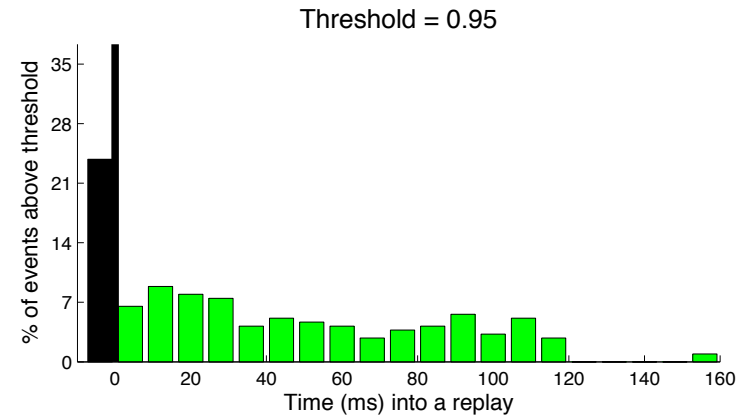
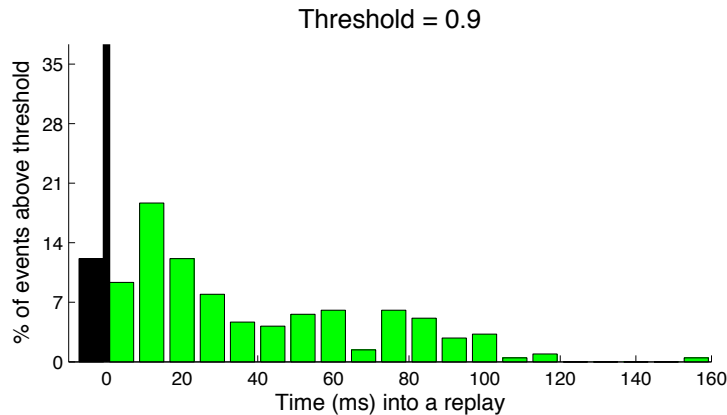
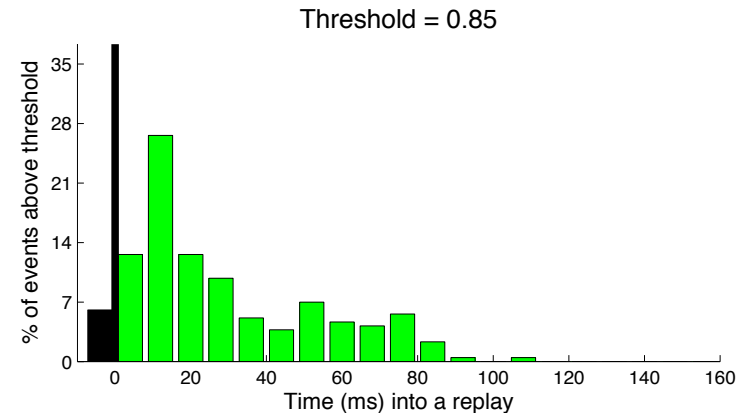
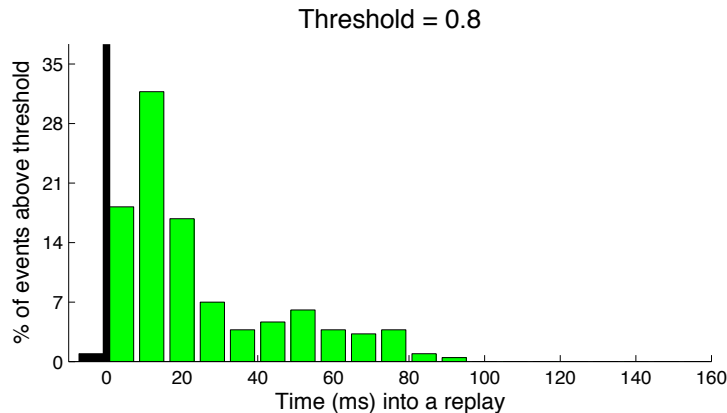


Replay 2



— outbound, forward — outbound, reverse — inbound, forward — inbound, reverse

Toward closed-loop interruption



- At a 0.8 confidence level, nearly all replay events can be classified within first 100 ms.

Applications of state space neural models

- Estimating ionic currents in dynamical neural models
- Identifying changing spiking patterns in Parkinson's disease
- Controlling neural motor prosthetic devices
- Tracking changing neural representations related to plasticity and learning
- Identifying switches between local representations of space and nonlocal representations.
- Transitions between tonic and bursting spiking regimes
- Identifying low dimensional manifolds of neural coactivation
- Selecting neurons that code for internal representations of value
- Real-time phase estimation for neural rhythms
- Joint decoding from spiking and LFP rhythms.

Conclusions

- Point process models can provide insight into how neural systems coordinate spiking activity to represent and process information.
- Marked point processes theory allows us to construct models of population coding, assess their ability to explain structure in spiking data, and make inferences about signals driving the neurons.
- Clusterless models allow for preservation of data, reduction of bias, and computationally efficient methods for online applications.

Acknowledgements

Xinyi Deng, BU

Long Tao, BU

Kensuke Arai, BU

Ali Yousefi, MGH

Felipe Gerhard, BU

Loren Frank, UCSF

Daniel Liu, UCSF

Kenneth Kay, UCSF

Eve Marder, Brandeis

Gabrielle Guitierrez, Brandeis

Tilman Kispersky, Brandeis

NSF IIS-0643995

NIH R01-NS073118

Simons Collaboration on the Global Brain