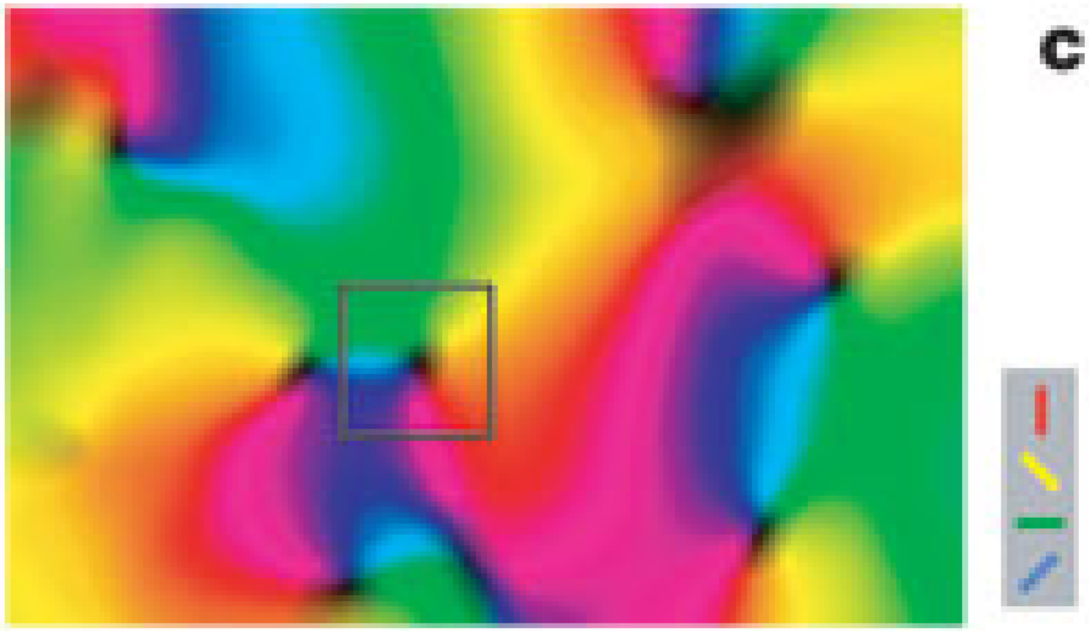
Efficient hierarchical receptive field estimation in simultaneously-recorded neural populations

Kamiar Rahnama Rad, Carl Smith, Timothy Machado, and Liam Paninski

Department of Statistics
Columbia University

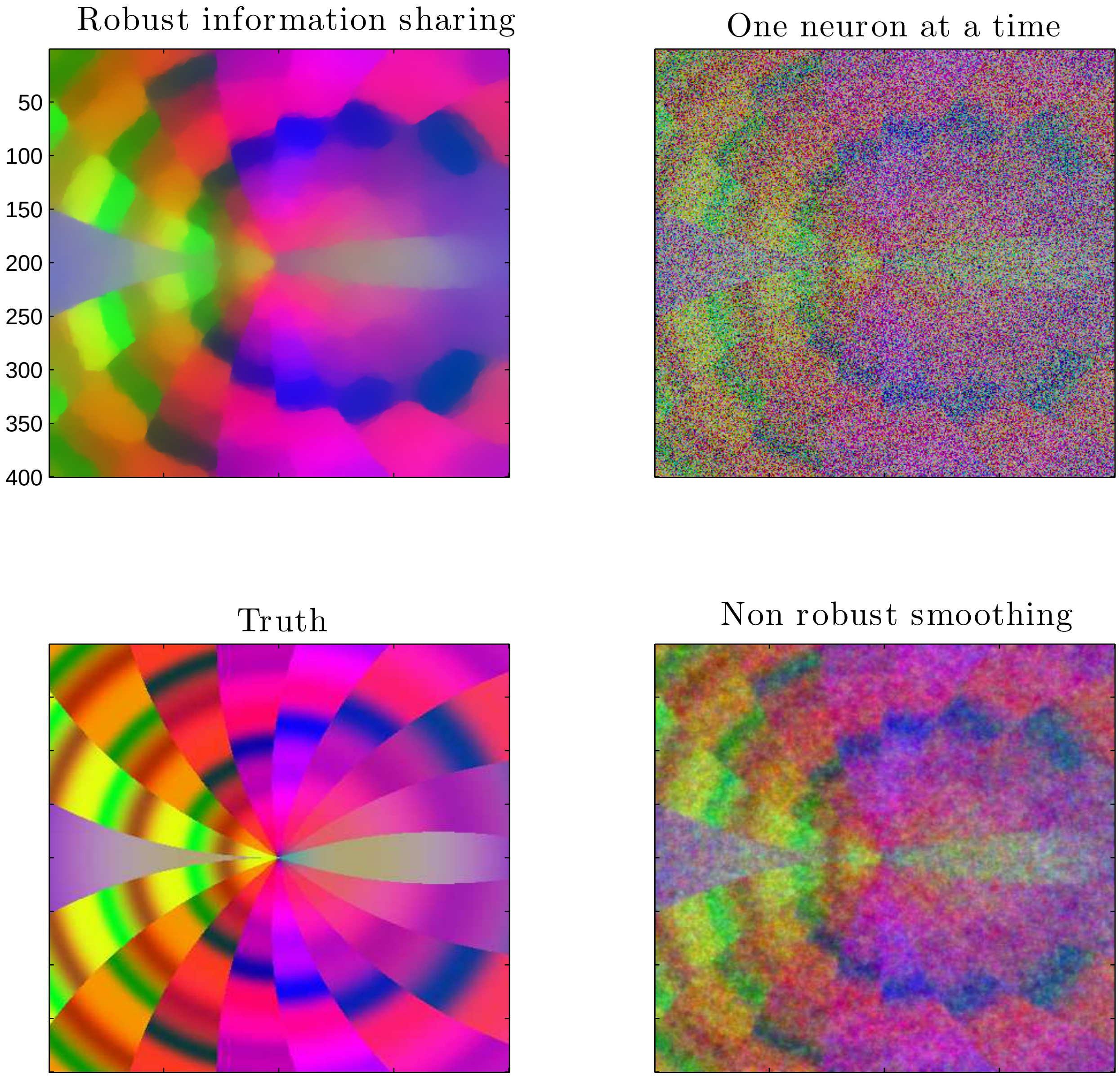
February 28, 2013

Representation of the visual environment in the brain



Orientation maps

- vary continuously across the cortical surface
- are punctuated by occasional jumps or discontinuities



- 160000 neurons
- 2 spikes per neuron

Neural encoding model

- Inhomogeneous point process with history dependence and coupling

$$
\begin{aligned}
r_{\ell,t} &\sim Poiss\left[\lambda_{\ell,t} dt\right] \\
\lambda_{\ell,t} &= f\left[(X_\ell \theta_\ell)_t + \sum_{\ell'} r_{\ell',t} * h_{\ell,\ell',t}\right]
\end{aligned}
$$

- θ_ℓ is the RF of neuron ℓ
- $(X_\ell \theta_\ell)_t$ is the projection of the stimulus onto the RF of neuron ℓ at time t

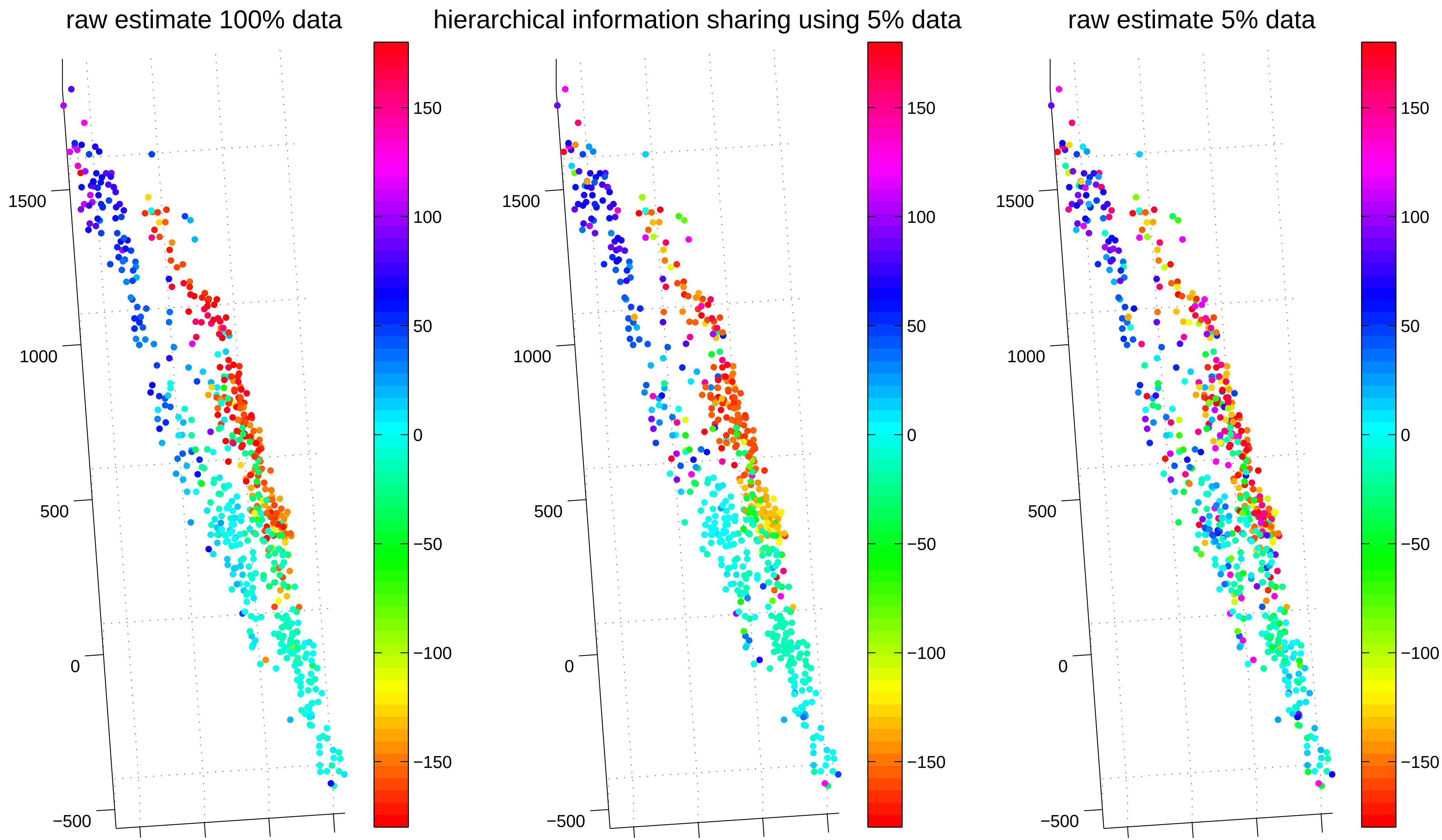
Hierarchical and robust joint estimation of the RF map

Maximum a posteriori estimate by maximizing

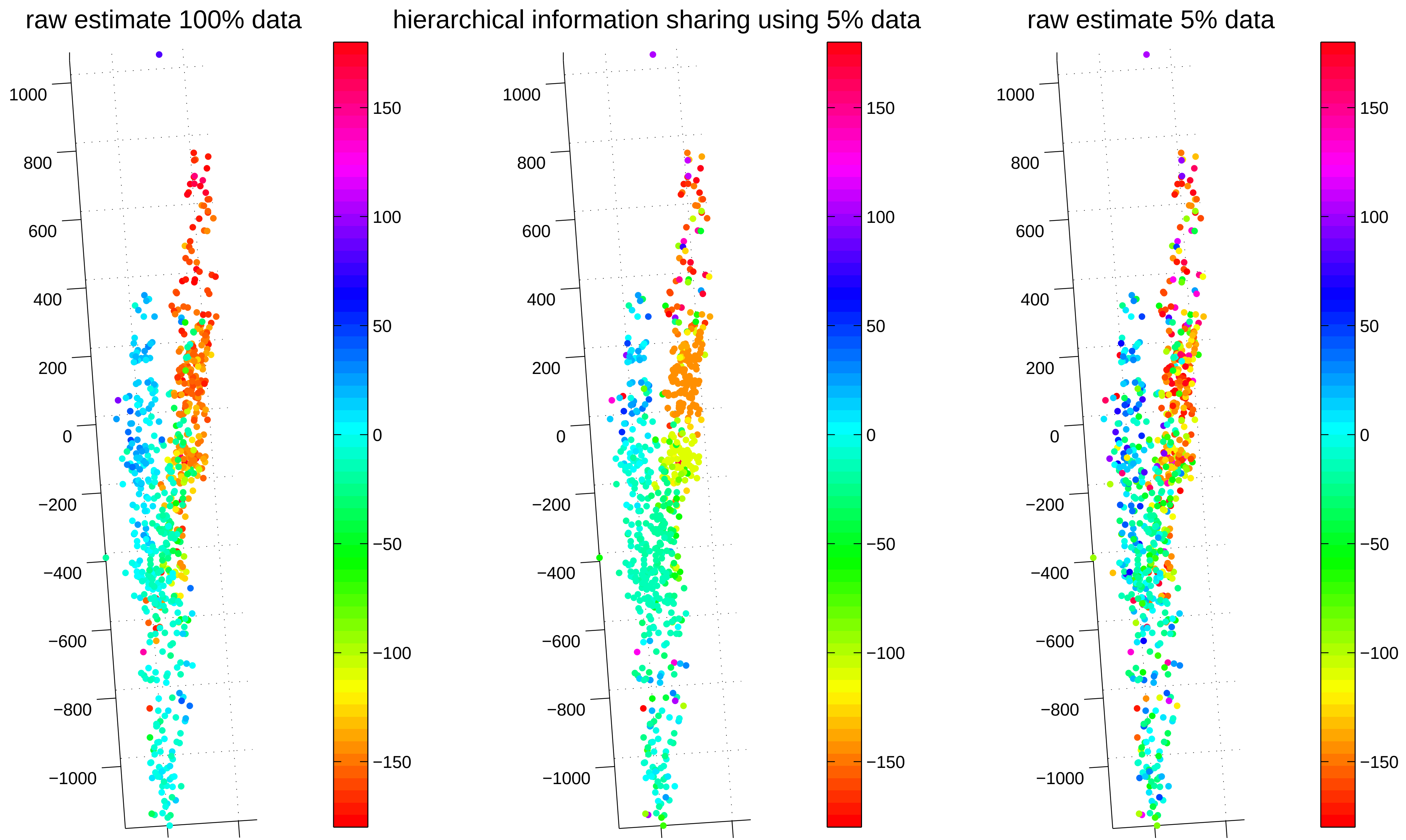
$$\log \Pr(\text{data}|\theta) - \lambda \sum_{\ell} \left\| D_\ell \theta \right\|_2 \quad D_\ell \theta = \begin{bmatrix} \theta_\ell - \theta_{\ell'} \\ \theta_\ell - \theta_{\ell_{\prime}} \end{bmatrix}$$

- neuron ℓ is located at (i, j)
- neuron ℓ' is located at $(i + 1, j)$
- neuron $\ell_{\prime}$ is located at $(i, j + 1)$

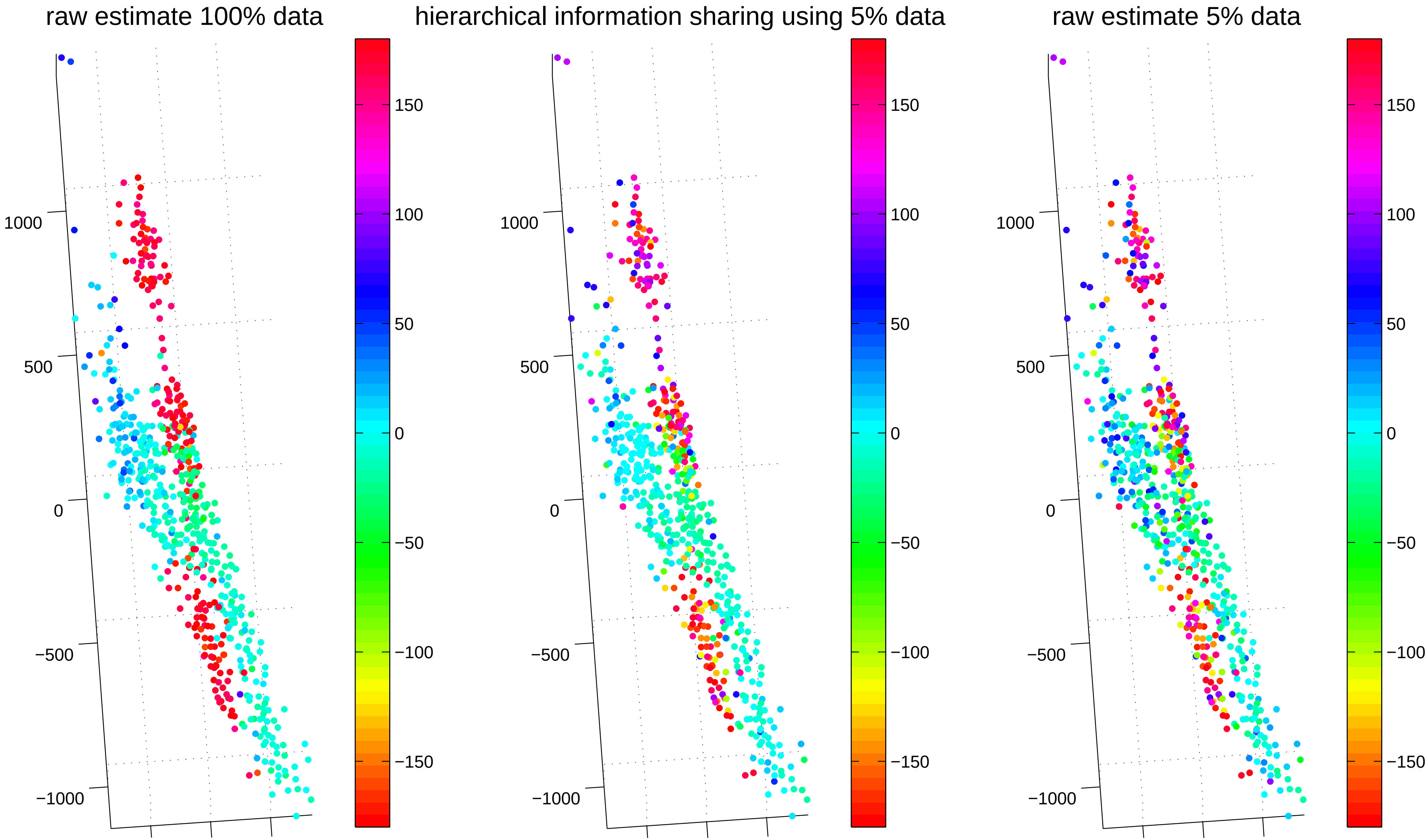
- Estimating receptive fields (or motor preferences) one neuron at a time is highly suboptimal.
- The precision of the border between functional maps can not be resolved unless the smoother is equipped with a right mix of prior information.



Phasic tuning at single-cell resolution



Phasic tuning at single-cell resolution



Phasic tuning at single-cell resolution

- $t_1, \cdots, t_L$ are spike times
- $\alpha_1 = \omega t_1, \cdots, \alpha_L = \omega t_L$ are noisy phases at spikes
- unknown phasic tuning ϕ
-

$$\begin{aligned} r \exp(i\alpha) &= \rho \exp(i\phi) + \text{noise} \\ \text{noise} &= \mathcal{N}\left(0, \sigma^2\right) + i\mathcal{N}\left(0, \sigma^2\right) \end{aligned}$$

Hierarchical and robust joint estimation of the RF map

- Maximum a posteriori estimate by maximizing

$$\log \Pr(\text{data}|\{\phi_\ell\}) - \lambda \sum_{\ell} \sum_{\ell' \in N_\ell} \left\| \begin{bmatrix} \rho_\ell \cos\phi_\ell - \rho_{\ell'} \cos\phi_{\ell'} \\ \rho_\ell \cos\phi_\ell - \rho_{\ell,} \cos\phi_{\ell,} \end{bmatrix} \right\|_2$$

- N_ℓ is the set of all neurons near to neuron ℓ

Mouse spinal cord

- Isolated neonatal mouse spinal cord contains neural circuits that can generate ordered patterns of periodic population activity
- Efficient characterization of the precise structure of phasic tuning at single-cell resolution
- Motor neuron activity was measured using large-scale, cellular resolution calcium imaging across hundreds of identified motor neurons
- See: II-63 Large-scale optical imaging reveals structured network output in isolated spinal cord. T. Machado, L. Paninski, T.M. Jessell

- Optimization has no non-global local minima
- Shares information across neurons: nearby neurons often have similar receptive fields
- Robust and adaptive: allows for large occasional breaks or outliers, contrary to previous work [1, 2].
- Posterior confidence intervals via Gibbs sampling, and scale mixtures
- Newton-Raphson iterations are fast: $O(d \log d)$ time; d = number of cells $\times$ dimensionality of each receptive field

Conlusion

- Estimating RFs one neuron at a time is highly inefficient
- Robust and adaptive information sharing can decreases the duration of the experiment up to 95%
- Adaptive experiment design can be done using posterior confidence intervals
- Hierarchical robust information sharing across neurons can scale to hundreds of thousands of simultaneously recorded neurons

Appendix

Probabilistic modeling

- Likelihood: $\ell = (i, j)$, $\theta_\ell \in \mathbb{R}^m$,

$$\begin{bmatrix} \vdots \\ r_\ell \\ \vdots \end{bmatrix} = \begin{bmatrix} \vdots \\ \theta_\ell \\ \vdots \end{bmatrix} + \sigma \begin{bmatrix} \vdots \\ \epsilon_\ell \\ \vdots \end{bmatrix}$$

- Prior:

$$\begin{aligned} \theta | \sigma^2, \tau_1^2, \cdots, \tau_{n^2}^2 &\sim \mathcal{N}(0, \sigma^2 C_\tau) \\ C_\tau^{-1} &= \sum_\ell \tau_\ell^{-2} D_\ell^T D_\ell = D^T \Gamma_\tau D \end{aligned}$$

$$\sigma^2, \tau_1^2, \cdots, \tau_{n^2}^2 \sim \pi(\sigma^2) d\sigma^2 \times \prod_{p=1}^{n^2} \frac{\lambda^2}{2} e^{-\lambda^2 \tau_p^2 / 2} d\tau_p^2 .$$

Posterior distribution via block Gibbs sampling

$$\theta|\sigma^2, \tau_1^2, \cdots, \tau_{n^2}^2, r \sim \mathcal{N}(\eta, C)$$

$$\eta = \left(I + D^T \Gamma_\tau D\right)^{-1} r$$

$$C = \sigma^2 \left(I + D^T \Gamma_\tau D\right)^{-1}$$

Posterior distribution via block Gibbs sampling

$$\sigma^2 | r, \theta, \tau_1^2, \cdots, \tau_{n^2}^2 \sim \pi(\sigma^2) d\sigma^2 \times \Gamma^{-1}(\alpha, \beta)$$

$$\Gamma^{-1}(\alpha, \beta) = \beta^\alpha x^{-\alpha-1} \frac{e^{-\beta/x}}{\Gamma(\alpha)}$$

$$\begin{aligned} \alpha &:= n^2 - 1, \\ \beta &:= \frac{1}{2}\|r - \theta\|^2 + \frac{1}{2}\theta^T C_\tau^{-1} \theta. \end{aligned}$$

Posterior distribution via block Gibbs sampling

$$\frac{1}{\tau_\ell^2}\Big| r, \theta, \sigma^2 \sim IG(\mu_\ell, \lambda_\ell).$$

where

$$\mu_\ell = \frac{\lambda\sigma}{\|D_\ell\theta\|_2}, \quad \lambda_\ell = \lambda^2,$$

$$IG(\mu, \lambda) = \sqrt{\frac{\lambda}{2\pi}} x^{-3/2} \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\}$$

Posterior distribution via block Gibbs sampling

$\theta|\sigma^2, \tau_1^2, \cdots, \tau_{n^2}^2, r \sim \mathcal{N}(\eta, C)$ where

$$\eta = \left(I + D^T \Gamma_\tau D\right)^{-1} r,$$
$$C = \sigma^2 \left(I + D^T \Gamma_\tau D\right)^{-1}$$

$$\sigma\left(I + D^T \Gamma_\tau D\right)^{-1}\left(\epsilon_1 + D^T \Gamma_\tau^{1/2} \epsilon_2\right) \sim \mathcal{N}(0, C)$$

References

[1] K. Rahnama Rad and L. Paninski. Efficient estimation of two-dimensional firing rate surfaces via gaussian process methods. *Network: Computation in Neural Systems*, 21:142–168, 2010.

[2] J.H. Macke, S. Gerwinn, L.E. White, M. Kaschube, and M. Bethge. Bayesian estimation of orientation preference maps. *NIPS*, 2010.