Efficient hierarchical receptive field estimation in simultaneously-recorded neural populations

Kamiar Rahnama Rad   Carl Smith   Liam Paninski
Columbia University

A major trend in systems neuroscience is to record simultaneously from large neuronal populations. A key objective in statistical neuroscience is to develop scalable and efficient methods for extracting as much information as possible from these recordings. One important direction involves hierarchical statistical modeling: estimating receptive fields (RFs) (or motor preferences) one neuron at a time is highly suboptimal, and in many cases we can do much better by sharing statistical information across neurons. In particular, we can exploit the fact that nearby neurons often have similar receptive fields. Here "nearby" might be defined topographically (e.g., in the case of cat primary visual cortex, where nearby neurons typically have similar orientation preferences) or more abstractly, in terms of, e.g., shared genetic markers.

We discuss two approaches for exploiting neighborhood information. The first method maximizes an appropriately penalized likelihood: we penalize deviations between neighboring RFs and compute the corresponding maximum a posteriori RF map. We use a smooth convex penalizer that allows for large occasional breaks or outliers in the inferred RF map. Posterior confidence intervals can be obtained here via "MAP-perturb" trick [1]. The second method is based on direct Gibbs sampling from the posterior, where the prior is of "low-rank" form, which enables fast direct sampling via the exact forward-backward approach discussed in [2]. Both approaches are computationally tractable, scalable to very large populations, and avoid imposing any overly restrictive constraints on the inferred RF map that would lead to oversmoothing. The first method is computationally cheaper, but the second method is able to model RFs in non-vector spaces (e.g., orientation). Both methods are equally applicable to multineuronal spike train or imaging data, and can dramatically reduce the experimental time required to characterize RF maps to the desired precision.

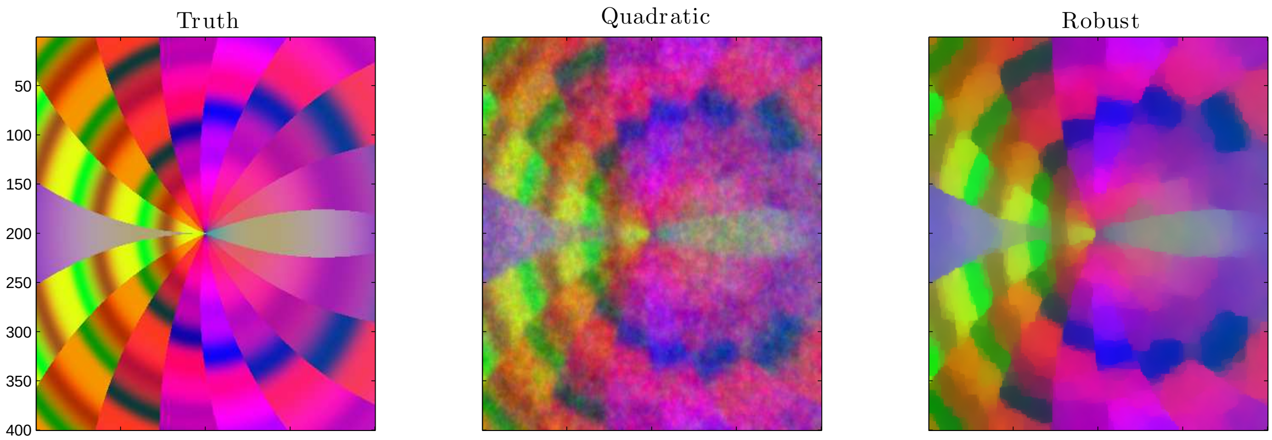
Supplementary material

We discuss the first approach described above in a bit more detail here, and provide an example illustrating the importance of hierarchical and robust joint estimation of the RF map. Let $\beta_p \in R^m$ be the RF of a neuron at location p, and let $\beta \in \mathbb{R}^{mn^2} = \left[\cdots, \beta_p^T, \cdots\right]^T$, assuming we are estimating neurons with m-dimensional receptive fields that live on an $n \times n$ spatial grid, for concreteness. (As emphasized above, this simple spatial setting can be generalized significantly.) In our first approach, the estimate β is chosen to minimize the objective function

$$-\log \Pr(\text{data}|\beta) + \tau \sum_p f\left(\left\| \begin{bmatrix} -\beta_{p''} + 2\beta_p - \beta_{p'} \\ -\beta_{p_{\prime\prime}} + 2\beta_p - \beta_{p_\prime} \end{bmatrix} \right\|_2 \right),$$

where $f(.)$ is a convex "Huber" function which behaves smoothly at zero but grows asymptotically linearly; $\beta_p, \beta_{p_\prime}, \beta_{p'} \in \mathbb{R}^m$ are the RFs at pixels (i,j), $(i,j+1)$ and $(i+1,j)$, respectively. The second term is the penalty that encourages nearby neurons to have similar inferred RFs.

This objective function is convex. As a result the optimization has no non-global local minima in β, and therefore standard ascent algorithms are guaranteed to converge to a global minimum. By exploiting the special structure of the problem, i.e., the sparsity of the Hessian, standard conjugate gradient minimization approaches can solve the problem very efficiently.



The figure on the left corresponds to an RF map of 160000 neurons; the RGB color indicates the three-dimensional vector β_p at each spatial location p. Our estimated map (based on the robust Huber prior) is shown on the right, and the estimated map based on a quadratic $f(.)$ — corresponding to a Gaussian process prior — is shown in the middle panel, for comparison. Borders between sharp breaks in the maps are preserved in the right panel, whereas the map in the middle is over smoothed.

References

G. Papandreou and A. Yuille. Gaussian sampling by local perturbations. In *NIPS*, pages 1858–1866, Vancouver, B.C., Canada, December 2010.

C. Smith, F. Wood, and L. Paninski. Low rank continuous-space graphical models. In *AISTAT*, 2012.