View Reviews

Paper ID
3060

Paper Title
Maximum-entropy and representative samples of neuronal activity: a dilemma

Reviewer #1

Questions

1. Please provide an "overall score" for this submission.
7: A good submission; an accept. I vote for accepting this submission, although I would not be upset if it were rejected.

2. Please provide a "confidence score" for your assessment of this submission.
5: You are absolutely certain about your assessment. You are very familiar with the related work.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.
The paper raises some interesting points about the use of maxent models to characterize statistical regularities in neural activity. It reviews some classic stats literature that is less familiar to a neuroscience audience and uses it to argue for a different way of estimating maxent models from neural data. Overall, it is a very nice paper; my only concern is that given the large review component and the target audience NIPS may not be the best venue for it.

Some highlights:
- what we are trying to estimate is statistical regularity patterns that expand to the large unobserved population. This kind if self-similarity constraints the form of the latents that influence the activity of the observed subpopulation.
- justification for using finite exchangeable priors for the prior distribution over patterns and nice mathematical properties guaranteeing that the observed neurons marginal is within the same class
- adding further constraints does not preserve this nice property - in general marginalizing out some dimensions in a maxent distribution does not result in a maxent distribution corresponding to the constraints on the remaining neurons.
- the joint estimate may be more meaningful given the kind of structure we are trying to extract from data.

Comments:
- would be interesting to link these observations more closely to recent attempts of using P(s) type of constraints to factor out shared latent structure - in the idea that this type of constraint would
- I think one should cite Tkacik et al 2012 as a first use of an finite exchangeable maxent distribution as a description on neural data
- the intro promised that the marginal vs joint fit comparison would provide some insight into some puzzling finds in the past literature; but I was missing that in the discussion
- I agree that it would be important to check to which extent the difference in estimated parameters would translate into qualitatively different conclusions in real data, especially after explicitly introducing P(s) type constraints.

Minor:

- using \Sigma to denote a random variable is a bit distracting

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?

2: Somewhat confident

Reviewer #2

Questions

1. Please provide an "overall score" for this submission.

5: Marginally below the acceptance threshold. I tend to vote for rejecting this submission, but accepting it would not be that bad.

2. Please provide a "confidence score" for your assessment of this submission.

3: You are fairly confident in your assessment. It is possible that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work. Math/other details were not carefully checked.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.

The authors suggest a method to account for the finite sample problem in maximum entropy estimations of underlying probability distributions. The approach here is to assume independent sampling from a population with a given population size (or one that comes from a known distribution). The sampling procedure is then marginalized over, giving rise to corrections to a “naïve” formula.

This is an important and extensively-studied problem. It seems that the specific solution proposed by the authors is new, although I am not an expert in this field.

But – because this is a highly studied topic, I think that the remark made in the discussion (Line 273-274) about comparing this to related frameworks cannot be left for future work. This is an essential part when suggesting a new method.

The paper has a very long introduction on the topic, and leaves little space for more results. For instance – plotting the error in certain observables as a function of sample size for the 3 methods of Figure 2, and a couple of methods from the literature.

Another aspect that is missing is an analysis of how the method is affected by properties of the underlying distributions. Do higher firing rates confer different errors than low firing rates? Do correlations affect the results?

Lines 276-279. It is widely acknowledged in neuroscience that we need to relate our knowledge of a population and its samples. This is discussed in many references in the paper itself, and more recent ones [1,2]. The Bayesian approach, which appears in many of these references, tries to explicitly state the assumptions (such as Equation 20 in the paper), and how they affect results.

[1] Wilting, J., & Priesemann, V. (2018). Inferring collective dynamical states from widely unobserved systems. Nature communications, 9(1), 2325.

[2] Gao, P., Trautmann, E., Byron, M. Y., Santhanam, G., Ryu, S., Shenoy, K., & Ganguli, S. (2017). A theory of multineuronal dimensionality, dynamics and measurement. bioRxiv, 214262.

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?

2: Somewhat confident

Reviewer #3

Questions

1. Please provide an "overall score" for this submission.
7: A good submission; an accept. I vote for accepting this submission, although I would not be upset if it were rejected.

2. Please provide a "confidence score" for your assessment of this submission.
3: You are fairly confident in your assessment. It is possible that you did not understand some parts of the submission or that you are unfamiliar with some pieces of related work. Math/other details were not carefully checked.

3. Please provide detailed comments that explain your "overall score" and "confidence score" for this submission. You should summarize the main ideas of the submission and relate these ideas to previous work at NIPS and in other archival conferences and journals. You should then summarize the strengths and weaknesses of the submission, focusing on each of the following four criteria: quality, clarity, originality, and significance.
The paper shows that the maximum entropy distributions obtained from using a fixed population of observed neurons vs. marginalizing a maximum entropy distribution from a larger population are different.

The paper looks sound and nicely points out a problem that people working with spike data and maximum entropy models might not be aware of. Many of the results are known in other fields and the authors are very clear in pointing that out. Previous literature is duely acknowledged.

My main issue with the paper is the question whether it matters in practice. Looking at the plots in Figure 2, the differences between the sample level and population level distribution seem minor. The sample level distribution even looks more "entropic" than the population level distribution since it puts more probability mass in the tail. In that sense the distribution deviates in the "right" direction w.r.t. maximum entropy. The paper would be much stronger if the authors could show a condition where the difference actually matter, for instance because one leads to wrong conclusions about cooperativity mentioned in the introduction. Even just pointing out the circumstances where the differences matter would be very useful. As it is, the paper rightfully states that there is a difference but leaves the reader wondering whether (s)he should worry about that in practice.

My second - minor - comment is that, even though the paper is generally well written and clear, it seems unncecessarily principled. There is nothing wrong with being principled per se, but I feel the paper could gain a lot in readability and smoothness if some of the more philosophical aspects would be omitted. They might be correct, but they break the flow when reading the paper.

4. How confident are you that this submission could be reproduced by others, assuming equal access to data and resources?
3: Very confident