Dear Editor,

We would like to submit our manuscript “Maximum-entropy and representative samples of neuronal activity: a dilemma” to Neural Computation.

The maximum-entropy approach to modelling correlated population activity in neural data, as well as other biological systems, has been of great interest in the past ten years. In a maximum entropy approach one typically tries to fit a distribution to the neural activity pattern where certain statistics (usually lower order statistics) are set to match the data while maximizing the entropy of the fit distribution.

An issue that has not been so far carefully addressed in the relevant studies is that the recorded population is a subsample of a larger population. This raises an important question, namely whether the maximum entropy approach should be directly applied to the sampled population, or whether one should estimate the statistics that would have been observed if the larger population was considered, fit the maximum entropy model and then marginalize to the subsample.

In this paper, we address the above question. Assuming a homogeneous maximum entropy approach where the statistics mean population activity, mean pairwise correlations etc, we construct the maximum entropy model along the two lines mentioned above. To peroform the second approach, namely including the large population, we use the observed statistics and use results from sampling theory to constrain the statistics of the activity of the larger population, and then fit a maximum entropy model.

Applied to neural data, we find that the two approaches lead to different distributions over the observed neurons. Since a population of recorded neurons is never isolated, the observed differences between the two approaches suggest that maximum entropy models fit directly to the subsampled population should be treated with caution and in general the large population approach should be preferred.

Best Regards,

P. Mana, V. Rostami, E. Torre and Y. Roudi