



Statistical analysis of multi-cell recordings: linking population coding models to experimental data

Jakob Macke^{1,2,3}, Philipp Berens^{1,2,4,5*} and Matthias Bethge^{1,2,4}

¹ Computational Vision and Neuroscience Group, Max Planck Institute for Biological Cybernetics, Tübingen, Germany

² Werner Reichardt Centre for Integrative Neuroscience and Institute of Theoretical Physics, University of Tübingen, Tübingen, Germany

³ Gatsby Unit for Computational Neuroscience, University College London, London, UK

⁴ Bernstein Center for Computational Neuroscience Tübingen, Tübingen, Germany

⁵ Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA

*Correspondence: philipp.berens@tuebingen.mpg.de

Modern recording techniques such as multi-electrode arrays and two-photon imaging methods are capable of simultaneously monitoring the activity of large neuronal ensembles at single cell resolution. These methods finally give us the means to address some of the most crucial questions in systems neuroscience: what are the dynamics of neural population activity? How do populations of neurons perform computations? What is the functional organization of neural ensembles?

While the wealth of new experimental data generated by these techniques provides exciting opportunities to test ideas about how neural ensembles operate, it also provides major challenges: multi-cell recordings necessarily yield data which is high-dimensional in nature. Understanding this kind of data requires powerful statistical techniques for capturing the structure of the neural population responses, as well as their relationship with external stimuli or behavioral observations. Furthermore, linking recorded neural population activity to the predictions of theoretical models of population coding has turned out not to be straightforward.

These challenges motivated us to organize a workshop at the 2009 Computational Neuroscience Meeting in Berlin to discuss these issues. In order to collect some of the recent progress in this field, and to foster discussion on the most important directions and most pressing questions, we issued a call for papers for this Research Topic. We asked authors to address the following four questions:

1. What classes of statistical methods are most useful for modeling population activity?
2. What are the main limitations of current approaches, and what can be done to overcome them?

3. How can statistical methods be used to empirically test existing models of (probabilistic) population coding?
4. What role can statistical methods play in formulating novel hypotheses about the principles of information processing in neural populations?

A total of 15 papers addressing questions related to these themes are now collected in this Research Topic. Three of these articles have resulted in “Focused reviews” in *Frontiers in Neuroscience* (Crumiller et al., 2011; Rosenbaum et al., 2011; Tchumatchenko et al., 2011), illustrating the great interest in the topic. Many of the articles are devoted to a better understanding of how correlations arise in neural circuits, and how they can be detected, modeled, and interpreted. For example, by modeling how pairwise correlations are transformed by spiking non-linearities in simple neural circuits, Tchumatchenko et al. (2010) show that pairwise correlation coefficients have to be interpreted with care, since their magnitude can depend strongly on the temporal statistics of their input-correlations. In a similar spirit, Rosenbaum et al. (2010) study how correlations can arise and accumulate in feed-forward circuits as a result of pooling of correlated inputs.

Lyamzin et al. (2010) and Krumin et al. (2010) present methods for simulating correlated population activity and extend previous work to more general settings. The method of Lyamzin et al. (2010) allows one to generate synthetic spike trains which match commonly reported statistical properties, such as time varying firing rates as well signal and noise correlations. The Hawkes framework presented by Krumin et al. (2010) allows one to fit models of recurrent population activity to the correlation-structure of experimental data. Louis et al. (2010) present a novel

method for generating surrogate spike trains which can be useful when trying to assess the significance and time-scale of correlations in neural spike trains. Finally, Pipa and Munk (2011) study spike synchronization in prefrontal cortex during working memory.

A number of studies are also devoted to advancing our methodological toolkit for analyzing various aspects of population activity (Gerwinn et al., 2010; Machens, 2010; Staude et al., 2010; Yu et al., 2010). For example, Gerwinn et al. (2010) explain how full probabilistic inference can be performed in the popular model class of generalized linear models (GLMs), and study the effect of using prior distributions on the parameters of the stimulus and coupling filters. Staude et al. (2010) extend a method for detecting higher-order correlations between neurons via population spike counts to non-stationary settings. Yu et al. (2010) describe a new technique for estimating the information rate of a population of neurons using frequency-domain methods. Machens (2010) introduces a novel extension of principal component analysis for separating the variability of a neural response into different sources.

Focusing less on the spike responses of neural populations but on aggregate signals of population activity, Boatman-Reich et al. (2010) and Hoerzer et al. (2010) describe methods for a quantitative analysis of field potential recordings. While Boatman-Reich et al. (2010) discuss a number of existing techniques in a unified framework and highlight the potential pitfalls associated with such approaches, Hoerzer et al. (2010) demonstrate how multivariate autoregressive models and the concept of Granger causality can be used to infer local functional connectivity in area V4 of behaving macaques.

A final group of studies is devoted to understanding experimental data in light of computational models (Galán et al., 2010; Pandarinath et al., 2010; Shteingart et al., 2010). Pandarinath et al. (2010) present a novel mechanism that may explain how neural networks in the retina switch from one state to another by a change in gap junction coupling, and conjecture that this mechanism might also be found in other neural circuits. Galán et al. (2010) present a model of how hypoxia may change the network structure in the respiratory networks in the brainstem, and analyze neural correlations in multi-electrode recordings in light of this model. Finally, Shteingart et al. (2010) show that the spontaneous activation sequences they find in cultured networks cannot be explained by Zipf's law, but rather require a wrestling model.

The papers of this Research Topic thus span a wide range of topics in the statistical modeling of multi-cell recordings. Together with other recent advances, they provide us with a useful toolkit to tackle the challenges presented by the vast amount of data collected with modern recording techniques. The impact of novel statistical methods on the field and their potential to generate scientific progress, however, depends critically on how readily they can be adopted and applied by laboratories and researchers working with experimental data. An important step toward this goal is to also publish computer code along with the articles (Barnes, 2010) as a successful implementation of advanced methods also relies on many details which are hard to communicate in the article itself. In this way it becomes much more likely that other researchers can actually use the methods, and unnecessary re-implementations can be avoided. Some of the papers in this Research Topic already follow this goal (Gerwinn et al., 2010; Louis et al., 2010; Lyamzin et al., 2010). We hope that this practice becomes more and more com-

mon in the future and encourage authors and editors of Research Topics to make as much code available as possible, ideally in a format that can be easily integrated with existing software sharing initiatives (Herz et al., 2008; Goldberg et al., 2009).

REFERENCES

- Barnes, N. (2010). Publish your computer code: it is good enough. *Nature* 467, 753.
- Boatman-Reich, D., Franaszczuk, P. J., Korzeniewska, A., Caffo, B., Ritzl, E. K., Colwell, S., and Crone, N. E. (2010). Quantifying auditory event-related responses in multichannel human intracranial recordings. *Front. Comput. Neurosci.* 4:4. doi: 10.3389/fncom.2010.00004
- Crumiller, M., Knight, B., Yu, Y., and Kaplan, E. (2011). Estimating the amount of information conveyed by a population of neurons. *Front. Neurosci.* 5:90. doi: 10.3389/fnins.2011.00090
- Galán, R. F., Dick, T. E., and Baekey, D. M. (2010). Analysis and modeling of ensemble recordings from respiratory pre-motor neurons indicate changes in functional network architecture after acute hypoxia. *Front. Comput. Neurosci.* 4:131. doi: 10.3389/fncom.2010.00131
- Gerwinn, S., Macke, J. H., and Bethge, M. (2010). Bayesian inference for generalized linear models for spiking neurons. *Front. Comput. Neurosci.* 4:12. doi: 10.3389/fncom.2010.00012
- Goldberg, D. H., Victor, J. D., Gardner, E. P., and Gardner, D. (2009). Spike train analysis toolkit: enabling wider application of information-theoretic techniques to neurophysiology. *Neuroinformatics* 7, 165–178.
- Herz, A. V. M., Meier, R., Nawrot, M. P., Schiegel, W., and Zito, T. (2008). G-Node: an integrated tool-sharing platform to support cellular and systems neurophysiology in the age of global neuroinformatics. *Neural Netw.* 21, 1070–1075.
- Hoerzer, G. M., Liebe, S., Schloegl, A., Logothetis, N. K., and Rainer, G. (2010). Directed coupling in local field potentials of macaque V4 during visual short-term memory revealed by multivariate autoregressive models. *Front. Comput. Neurosci.* 4:14. doi: 10.3389/fncom.2010.00014
- Krumin, M., Reutsky, I., and Shoham, S. (2010). Correlation-based analysis and generation of multiple spike trains using Hawkes models with an exogenous input. *Front. Comput. Neurosci.* 4:147. doi: 10.3389/fncom.2010.00147
- Louis, S., Gerstein, G. L., Grün, S., and Diesmann, M. (2010). Surrogate spike train generation through dithering in operational time. *Front. Comput. Neurosci.* 4:127. doi: 10.3389/fncom.2010.00127
- Lyamzin, D. R., Macke, J. H., and Lesica, N. A. (2010). Modeling population spike trains with specified time-varying spike rates, trial-to-trial variability, and pairwise signal and noise correlations. *Front. Comput. Neurosci.* 4:144. doi: 10.3389/fncom.2010.00144
- Machens, C. K. (2010). Demixing population activity in higher cortical areas. *Front. Comput. Neurosci.* 4:126. doi: 10.3389/fncom.2010.00126
- Pandarinath, C., Bomash, I., Victor, J. D., Prusky, G. T., Tschetter, W. W., and Nirenberg, S. (2010). A novel mechanism for switching a neural system from one state to another. *Front. Comput. Neurosci.* 4:2. doi: 10.3389/fncom.2010.00002
- Pipa, G., and Munk, M. H. J. (2011). Higher order spike synchrony in prefrontal cortex during visual memory. *Front. Comput. Neurosci.* 5:23. doi: 10.3389/fncom.2011.00023
- Rosenbaum, R., Trousdale, J., and Josi, K. (2011). The effects of pooling on spike train correlations. *Front. Neurosci.* 5:58. doi: 10.3389/fnins.2011.00058
- Rosenbaum, R. J., Trousdale, J., and Josi, K. (2010). Pooling and correlated neural activity. *Front. Comput. Neurosci.* 4:9. doi: 10.3389/fncom.2010.00009
- Shteingart, H., Raichman, N., Baruchi, I., and Ben-Jacob, E. (2010). Wrestling model of the repertoire of activity propagation modes in quadruple neural networks. *Front. Comput. Neurosci.* 4:25. doi: 10.3389/fncom.2010.00025
- Stauder, B., Grün, S., and Rotter, S. (2010). Higher-order correlations in non-stationary parallel spike trains: statistical modeling and inference. *Front. Comput. Neurosci.* 4:16. doi: 10.3389/fncom.2010.00016
- Tchumatchenko, T., Geisel, T., Volgushev, M., and Wolf, F. (2010). Signatures of synchrony in pairwise count correlations. *Front. Comput. Neurosci.* 4:1. doi: 10.3389/fncom.2010.00010
- Tchumatchenko, T., Geisel, T., Volgushev, M., and Wolf, F. (2011). Spike correlations – what can they tell about synchrony? *Front. Neurosci.* 5:68. doi: 10.3389/fnins.2011.00068
- Yu, Y., Crumiller, M., Knight, B., and Kaplan, E. (2010). Estimating the amount of information carried by a neuronal population. *Front. Comput. Neurosci.* 4:10. doi: 10.3389/fncom.2010.00010

Received: 15 June 2011; accepted: 14 July 2011; published online: 28 July 2011.

Citation: Macke J, Berens P and Bethge M (2011) Statistical analysis of multi-cell recordings: linking population coding models to experimental data. *Front. Comput. Neurosci.* 5:35. doi: 10.3389/fncom.2011.00035

Copyright © 2011 Macke, Berens and Bethge. This is an open-access article subject to a non-exclusive license between the authors and Frontiers Media SA, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and other Frontiers conditions are complied with.