

Decoding the activity of neural networks: biological decoding and constant metric content

Máté Lengyel^{1,2*}, Péter Érdi^{1,3}, Alessandro Treves²

¹Computational Neuroscience Group, Dept. Biophysics, KFKI Research Institute for Particle and Nuclear Physics, Hungarian Academy of Sciences, Budapest, Hungary, ²Neuroscience Sector, SISSA International School for Advanced Studies, Trieste, Italy, ³Center for Complex Systems Studies, Kalamazoo College, Kalamazoo, MI, USA
*corresponding author: Máté Lengyel, KFKI R.I.P.N.P., Hun Acad Sci, 29-33 Konkoly-Thege M. út, Budapest H-1121, Hungary

A decoding algorithm is a procedure that assigns to neuronal activity a prediction or probabilistic guess about what in the external world (which sensory stimuli, or else motor outputs, or which state of the animal) the activity is coding for. Thus decoding is the inverse of coding, and if done correctly it allows to estimate how informative coding was in the first place. This is because the amount of information that the decoded stimuli bear about the actual stimuli can never exceed the amount of information in the activity itself, and it approaches it if decoding is efficient. Neurons in the real nervous system can also in some cases be thought to perform decoding of their inputs. This is when an ensemble of neurons represent a set of sensory stimuli using some population code, and neurons in downstream areas “read out” the original stimuli off the activity of their afferents, in order to carry out their own information processing. From a technical point of view, decoding algorithms can also be useful as possible solutions to the limited sampling problem, in particular when recording responses from behaving animals in which sampling is perforce limited [1]. Two problems related to how neurons decode sensory stimuli were addressed.

I. Decoding can be done in at least two markedly different ways [1]. First, one can try to optimize the decoding algorithm, to minimize the loss of information it entails. Thus a Bayesian decoding scheme results in values of mutual information (MI) and fraction of correct decodings (f_{corr}) that can be taken as the theoretical limit any decoding can reach. Bayesian decoding usually requires an explicit model of the distribution of neuronal responses, however, which can be difficult to obtain, in practice. More importantly, the mathematical operations involved are hardly carried out by real neurons. Second, one can try to imitate the operations carried out by real neurons. A biologically more realistic decoding method, the dot-product (DP) algorithm, has been widely used, but it can in many cases lead to substantially sub-optimal decoding results, i.e. to much lower information values [2].

By mathematical analysis we investigated conditions which tend to make the DP method provide equivalent decoding as a Bayesian algorithm. We determined a set of three constraints. (1) The general form of response distributions had to be asymmetric, of high variance and greatly overlapping between stimuli. (2) The length of synaptic weight vectors of neurons had to be normalized in a non-trivial way: weight vectors for decoding stimuli resulting in heterogeneous neuronal activity (some neurons are being much more active than others) had to be longer than those decoding homogeneous activity. (3) The direction of synaptic weight vectors in most cases needed to strongly deviate from that of the average of response vectors corresponding to the respective stimuli. The latter two constraints conspicuously violate the Hebbian learning paradigm, and make it unlikely that our response model could be produced by Hebbian plasticity. We propose possible alternative learning rules that may provide neurons with the necessary weight vectors for achieving quasi-optimal decoding.

II. Metric content (mc), defined from MI and f_{corr} , describes how much metricity is recognized in a set of stimuli by a group of neurons: it is zero if all incorrect decodings are equiprobable, and thus equidistant from the correct one; and it is one if incorrect decodings are perfectly clustered into two groups, one including those being equiprobable with the correctly decoded stimulus, and one including those having zero probability [3]. It had been shown by previous analysis of experimental data collected in monkeys and rats that while both MI and f_{corr} increase as more neurons are used for decoding, mc tends to remain constant, as long as only neuronal responses recorded from the same brain region are used for decoding [2]. Different brain regions are characterized by different values of mc [3].

In numerical Monte-Carlo simulations we constructed a simple response model: a neuron produced all possible responses but one with equal probability for a given stimulus. Tuning precision (tp), being defined as the difference between the probability of the preferred and that of the other responses ($tp=0$ indicates all responses being equiprobable, and $tp=1$ that only one response is possible per stimulus), was kept constant across all neurons, and the preferred stimulus was randomly selected for each stimulus and for each neuron. Two variants were compared for the calculation of information measures: the confusion matrix for 'real' and 'predicted' stimuli was either directly calculated from the response probability tables, as would be the case in an ideal situation with unlimited sampling, or, to better approach an experimental situation with the limited sampling problem, indirectly by first simulating 1000 trials for each stimulus and then constructing confusion matrices from the 'collected' data.

Our model reproduced the increase in both MI and f_{corr} as the number of neurons used for decoding increased. More specifically, MI exponentially saturated at the entropy of the stimulus set, as predicted by a simple model [1,4], and as found, roughly, in experiments [2,5]. Higher values of tp resulted in higher MI , f_{corr} and mc values. We found that both MI and f_{corr} were overestimated in the finite sample case at low values of tp ($tp < 0.8$), but reliably matched values extracted from the ideal case at $tp > 0.8$. Conversely, mc could be estimated more accurately at low tp values. More importantly, a strong correlation between the tp and mc of the population code was demonstrated: as long as tp was kept constant across neurons, mc also remained nearly uniform. However, as tp grew, deviations from the average mc became larger, rising from ~ 0.01 to $\sim 0.2-0.3$ in units of average mc , thus conflicting with the nearly constant values noted in experimental findings. When noise and signal correlations in the population response were also taken into account in the construction of the response model, these deviations were decreased. This indicates that the metric content of neural population responses may be strongly influenced by neural oscillations, known to increase noise correlations between neurons, and by topographic feed-forward and recurrent connections, known to be a major source of signal correlations.

- [1] Rolls ET, Treves A. *Neural Networks and Brain Function*. Oxford University Press, 1998.
- [2] Rolls ET, Treves A, Tovee MJ. The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. *Exp Brain Res* 114:149-162, 1997.
- [3] Treves A. Information coding in higher sensory and memory areas, *In: Handbook of Biological Physics* (eds. F Moss, S Gielen), vol.4, pp. 803-829, 2000.
- [4] Samengo I, Treves A. Representational capacity of a set of independent neurons. *Phys Rev E* 63:011910, 2001.
- [5] Gawne TJ, Richmond BJ. How independent are the messages carried by adjacent inferior temporal cortical neurons? *J Neurosci* 13:2758-71, 1997.