‘Dynamics of neuronal interactions’ cannot be explained by ‘neuronal transients’

E. VAADIA\(^{1}\), A. AERTSEN\(^{2}\) and I. NELKEN\(^{3}\)

\(^{1}\) Department of Physiology, Hadassah School of Medicine & The Center for Neural Computation, The Hebrew University, Jerusalem, 91120 Israel
\(^{2}\) The Center for Research of Higher Brain Functions, Department of Neurobiology, The Weizmann Institute of Science, Rehovot, 76100 Israel
\(^{3}\) Department of Physiology, Hadassah School of Medicine, The Hebrew University, Jerusalem, 91120 Israel

SUMMARY

In a recent paper, Vaadia \textit{et al.} demonstrated that patterns of firing correlation between single neurons in the cortex of behaving monkeys can be modified within a fraction of a second. These changes occur in relation to sensory stimuli and behavioral events, and even without modulations of the neurons’ firing rates. These findings call for a revision of prevailing models of neural coding that solely rely on single neuron firing rates. In a defense of these models, Friston put forward an alternative explanation, proposing that the observed correlation dynamics emerge solely from co-modulations of the firing rates of each of the neurons, while the strength of their interaction remains constant. To test this possibility we re-examined the data, adopting Friston’s ‘neuronal transients’ model, and the associated equations and procedures. We found that, to explain the dynamic correlation between a pair of neurons, the alternative interpretation requires that each neuron’s response to a single stimulus is composed of a relatively large number of independent components, which co-vary with their counterparts in the companion neuron. This large number of components and their shapes lead us to conclude that, although in principle possible, the neuronal transients model: (i) does not provide a simpler explanation of the experimental results; and (ii) cannot explain these results without itself deviating significantly from most rate code models.

In his paper ‘Neuronal transients’ Friston (1995) presents an alternative explanation of the results of the joint peri-stimulus time (jprt) analysis for the spike trains of pairs of simultaneously recorded frontal cortex neurons by Vaadia \textit{et al.} (1995). One of the prime goals of his explanation is to ‘reinstate’ prevailing models of neuronal coding based on firing rates. To test if the neuronal transient model provides an adequate explanation for our experimental findings, we re-analysed the data we presented by using the analysis and equations (1), (2) and (4) in Friston’s paper (Friston 1995). First, we estimated the ‘non-specific correlation’. Then, by using singular value decomposition (svd) analysis (which is a generalization of principal components analysis to non-symmetric and not-square matrices; for more information see Press \textit{et al.} 1988), we computed the hypothetical ‘neuronal transients’ that best explain the data. Finally, we computed the correlation matrix, generated by the weighted contributions of the ‘neuronal transients’ and the ‘non-specific correlation’ (see legend of figure 1 for details of this inverse approach). We found that the neuronal transients model does not provide a simple explanation of our results.

The results of the analysis for the spike trains of one neuronal pair from figure 2a of Vaadia \textit{et al.} (1995) are compared to the results of jprt analysis in figures 1–3. We first examined the ‘non-specific’ (constant) correlation. By its stationary nature, it cannot account for dynamic modulation of correlation along the diagonal. Rather, it changes the excess correlation along the diagonal by a constant amount, without affecting its time course. This is illustrated in figure 1 for the example discussed here. The solid curve replicates the coincidence-time histogram from our jprt-matrix (the green histogram in Vaadia \textit{et al.} (1995), see figure 2a). The correlation gradually builds up, starting 1000 ms before the stimulus, and reaches a peak about 400 ms after it. After subtraction of the estimated ‘non-specific’ correlation (dotted line), the correlation dynamics maintain a virtually identical time course (dashed line). Thus, Friston’s proposition that non-specific correlation would account for ‘marked correlation before the stimulus arrives’ is not substantiated. We note in passing that event-related ‘anticipatory’ activity (be it rate or correlation) preceding stimuli or other behavioral events is frequently encountered in the frontal cortex of behaving monkeys (Tanji & Evarts 1976; Vaadia \textit{et al.} 1988).

To evaluate the combined contributions of non-specific correlation and co-varying ‘neuronal transi-
Figure 1. Dynamic modulation of correlated firing. The solid line replicates the diagonal (green) coincidence-time histogram in Vaadia et al. (1995) figure 2a. The dotted line is the estimated ‘non-specific correlation’ for the same data. The difference between these two graphs is shown in the dashed line. The small fluctuations of the ‘non-specific’ correlations are the result of fluctuations in the estimated standard deviation (the non-specific covariance is constant, but we divided it by the standard deviation to obtain the non-specific correlation as follows. We re-analysed the data in Vaadia et al. (1995), figure 2a, using svd-analysis and the model equations (1), (2) and (4) in Friston (1995). Since, unlike Friston, because we based our analysis on physiological data, the values of the ‘non-specific activities’ \( \xi \) and \( \eta \) were not available. Therefore the non-specific covariance \( \gamma_{ij} \) was estimated by computing the covariance between the two spike trains over the time window between 1500-2500 ms after the ‘ready’ signal (cf. figure 3a). During this interval there is no discernible modulation in the \( \text{psr} \) correlation matrix. This estimate, extended along the entire diagonal, was subtracted from the covariance matrix (the raw \( \text{psr} \) minus the \( \text{psr} \)-predictor \( \langle \text{Aertsen et al. 1989} \rangle \)), and the resulting difference matrix was subjected to svd-analysis. The first pair of singular vectors (see figure 2, top traces) were taken as estimators of the hypothetical ‘rate transients’ \( \tau_r \) and \( \tau_j \). To produce the prediction of the svd-analysis for our normalized correlation matrix, we first evaluated the contribution of the first term in equation (2) of Friston (1995) (the product of the first singular vectors, multiplied by the corresponding singular value). We then added back the non-specific covariance \( \gamma_{ij} \). Finally, we divided the resulting covariance matrix by the standard deviation matrix \( \langle \text{Aertsen et al. 1989} \rangle \) and obtained the predicted correlation matrix (figure 3b). We note that there are practical difficulties with estimating the ‘non-specific correlation’ from spike train data. Experimental observations show different amounts of correlation under different conditions and at different times, even before or in the absence of a stimulus. This is also the case in the present example. Had we, for example, the 1000 ms preceding the ready signal in figure 3a to estimate \( \gamma_{ij} \), we would have arrived at a higher non-specific correlation. This, however, would mainly shift the dashed line downwards, and hardly affect its time course.)

Figure 2. The first five pairs of singular vectors computed by svd-analysis of the spike train data presented in figure 2a of Vaadia et al. (1995). The corresponding singular values decay slowly \( (\lambda_1-\lambda_5; 256, 237, 200, 169, 148) \). Accordingly, there is only a slow increase in the variance accounted for by the first few singular vectors (the fraction of the variance contributed by each of the first five vectors are: 17%, 14%, 10%, 6%, 6%).

find it hard to associate them with any consistent modulation of the firing rates, let alone to a physiological process that would be linked with them.

Thus, our analysis indicates that the neuronal transients model can explain the experimental data only if four conditions are met: (i) the responses of the two neurons $i$ and $j$ to a specific stimulus are characterized by 5–10 different ‘transients’ (firing rate modulations $\tau_{ik}^s$ and $\tau_{jk}^s$, $k = 1–10$); (ii) each stimulus presentation evokes a mixture of these transients with random weight factors $\alpha_k^i$ and $\alpha_k^j$; (iii) the weight factors of the two neurons are uncorrelated for different $k$’s; but, (iv) they are correlated for the same $k$ (the singular value $\lambda_k$ is a measure of this correlation). We conclude that Friston’s alternative interpretation is, in principle, possible but is not ‘a simpler explanation’ (Friston 1995) of our results.

Under these conditions, and particularly in view of the need to invoke relatively large numbers of transients, the neuronal transients model deviates significantly from most rate code models. Also, preliminary svd-analysis of pair correlations among a simultaneously recorded triplet of neurons suggests that the dominant singular vectors of a given neuron may be different, depending on the selected ‘partner’ neuron. This would imply that the ‘transients’ are not only a property of the single neuron, but also of the selected pair. Evidently, this is not what could be expected from a rate code model.

Nevertheless, we do believe that Friston makes two important points, that provide valuable suggestions for future studies.

1. svd-analysis is certainly a helpful tool for the analysis of neuronal activity. It could be used to evaluate typical temporal patterns of firing rate modulations, and to characterize and quantify the variability of these modulations across repeated presentations of the same stimulus. A combination of svd-analysis with measurements of spike correlation can be used to assess the relative contributions of these modulations to the neural code.

2. The incorporation of ‘nonspecific correlation’ provides an attractive alternative null hypothesis for the study of neuronal interactions. Rather than to test against the classical null hypothesis of zero correlation (independent firing) as the current jprst-analysis does, it might indeed be more appropriate (stronger) to test for the presence of correlation dynamics against the null hypothesis of constant correlation, provided a good experimental estimate of this constant correlation can be obtained.

In spite of this, however, the prime objective of Friston’s proposal is not met. The potential attraction of his model would be that it might explain our experimental results without the need to go beyond the widely accepted firing rate models. As we have demonstrated, however, this appears to be impossible, without deviating from the conventional rate model. We maintain that a more attractive alternative is, in fact, to venture beyond single neuron rate coding and to incorporate the correlation of spike firing as a promising new mechanism to dynamically organize neurons into functional groups.
We thank M. Abeles, H. Bergman, I. Haalman, Y. Prut and H. Slovin for their contributions. The study was supported in part by grants from the Israeli Academy of Sciences and Humanities, the American-Israel Bi-national Science Foundation (BSF), and the ‘Human Frontier Science Program (HFSP)’.

REFERENCES


Received 2 June 1995; accepted 20 June 1995
Figure 3. (a) Normalized pST correlation matrix from figure 2a of Vaadia et al. (1995). (b) Estimation of the correlation matrix in (a), based on the combined contributions of non-specific correlation and the first pair of singular vectors weighted by the corresponding singular value (constructed as described in the legend of figure 1).