

This is an appendix to the paper by Panzeri et al 1999 Correlations and the encoding of information in the nervous system. Proc. R. Soc. Lond. B **266**, 1001-1012.

Electronic appendices are refereed with the paper. However, no attempt has been made to impose a uniform editorial style on the electronic appendix.

Appendix B: Evaluation of the bias and of the variance of information derivatives

It is possible to analytically derive an estimate of the amount of the bias, which can then be subtracted to provide an unbiased estimate. This is done using the standard error propagation procedure (see e.g. Bevington and Robinson, 1992).

A function of the firing rates can be expanded about the mean rate as

$$\langle f(\bar{r}_i) \rangle \approx f(\bar{r}_i) + \frac{1}{2} \sum_i \frac{\partial^2 f}{\partial \bar{r}_i^2} \sigma_{\bar{r}_i}^2 \quad (\text{B.1})$$

where

$$\sigma_{\bar{r}_i}^2 = \frac{1}{N-1} (\overline{r_i^2} - \bar{r}_i^2). \quad (\text{B.2})$$

Applying this to Eq. 8, we obtain

$$I_{t(\text{bias})} = \frac{1}{2 \ln 2} \sum_s \widehat{\sum}_i \left(\frac{P(s)}{\bar{r}_i(s)} - \frac{P(s)^2}{\bar{r}_i} \right) \sigma_{\bar{r}_i(s)}^2 \quad (\text{bits/sec}) \quad (\text{B.3})$$

where the ‘hat’ over the i summation indicates that it is only over the ‘relevant’ s, i pairs, i.e. those with non-zero underlying probability of spike emission. If the underlying probability is zero, then no finite sampling fluctuations are possible and that s, i does not contribute to the bias.

Obviously this correction as it stands cannot be applied to the components of I_{tt} . A similar correction can be derived by the same method. This calculation as we shall see is slightly more involved. We will have to calculate the bias for each component of I_{tt} separately, so we will consider a generic function $f(\bar{\mathbf{x}})$ of a set $\{\bar{x}_j\}$ of (possibly correlated) random variables. Each random variable \bar{x}_j is the average (obtained on the basis of a limited number of trials N) of a random variable x_j . We assume for the purposes of our analytical estimate that the number of trials N is large but finite. In this case the independent variables \bar{x}_j fluctuate around their true value $\langle \bar{x}_j \rangle$, and the fluctuations scale as $1/N$. Therefore this derivation of the bias of each information component using error propagation is equivalent to the $1/N$ expansion of the bias of the full information derived e.g. in (Panzeri and Treves, 1996).

Under these assumptions, the sampling bias in $f(\bar{\mathbf{x}})$ is:

$$\langle f(\bar{\mathbf{x}}) \rangle - f(\langle \bar{\mathbf{x}} \rangle) \simeq \frac{1}{2} \sum_j \frac{\partial^2 f}{\partial \bar{x}_j^2} (\langle \bar{\mathbf{x}} \rangle) \sigma_{\bar{x}_j}^2 + \sum_{i < j} \frac{\partial^2 f}{\partial \bar{x}_i \partial \bar{x}_j} (\langle \bar{\mathbf{x}} \rangle) \sigma_{\bar{x}_i \bar{x}_j}^2, \quad (\text{B.4})$$

where the (co)variances of the means of the random variables $\{x_j\}$ are

$$\sigma_{\bar{x}_j}^2 = \frac{1}{N-1} \left(\langle \bar{x}_j^2 \rangle - \langle \bar{x}_j \rangle^2 \right) \quad (\text{B.5})$$

$$\sigma_{\bar{x}_i \bar{x}_j}^2 = \frac{1}{N-1} \left(\langle \bar{x}_i \bar{x}_j \rangle - \langle \bar{x}_i \rangle \langle \bar{x}_j \rangle \right) \quad (\text{B.6})$$

The expression for the variance of $f(\bar{\mathbf{x}})$ is similarly

$$\sigma_{f(\bar{\mathbf{x}})}^2 = \sum_j \left(\frac{\partial f}{\partial \bar{x}_j} \right)^2 \sigma_{\bar{x}_j}^2 + 2 \sum_{j < i} \left(\frac{\partial f}{\partial \bar{x}_j} \right) \left(\frac{\partial f}{\partial \bar{x}_i} \right) \sigma_{\bar{x}_j \bar{x}_i}^2 \quad (\text{B.7})$$

To apply the above formalism to our cases of the information derivative components $I_{tt1} - I_{tt3}$, the random variables used are the $S \times C$ mean rates of the cells to each stimulus $\bar{r}_i(s)$, and the $S \times C(C + 1)/2$ variables $\bar{\kappa}_{ij}(s)$, which are defined as

$$\bar{\kappa}_{ii}(s) = \bar{r}_i^2(s)(1 + \gamma_{ii}(s)) = \frac{\overline{n_i^2(s) - n_i(s)}}{t^2}, \quad i = 1, \dots, C \quad (\text{B.8})$$

$$\bar{\kappa}_{ij}(s) = \bar{r}_i(s)\bar{r}_j(s)(1 + \gamma_{ij}(s)) = \frac{\overline{n_i(s)n_j(s)}}{t^2}, \quad i, j = 1, \dots, C; i < j. \quad (\text{B.9})$$

$\bar{\kappa}_{ij}(s) (i \neq j)$ represents the rate of coincidences between cells, whereas $\bar{\kappa}_{ii}(s)$ parameterises the rate of ‘autoincident’ firing by the same cell (i.e. two different spikes fired by the one cell in the same small time window). The corresponding $\gamma_{ij}(s)$ parameters could have been used in this derivation, but use of the $\bar{\kappa}$ variables in this section makes the bias derivation easier and more transparent. These random variables are each calculated on the basis of N_s trials per stimulus.

The leading contributions of these (co)variances can be calculated analytically in the short time window limit. They are:

$$\sigma_{\bar{r}_i(s)}^2 \simeq \frac{\bar{r}_i(s)}{tN_s} \quad (\text{B.10})$$

$$\sigma_{\bar{\kappa}_{ii}(s)}^2 \simeq \frac{2\bar{\kappa}_{ii}(s)}{t^2N_s} \quad (\text{B.11})$$

$$\sigma_{\bar{\kappa}_{ij}(s)}^2 \simeq \frac{\bar{\kappa}_{ij}(s)}{t^2N_s} \quad (i \neq j) \quad (\text{B.12})$$

$$\sigma_{\bar{r}_i(s)\bar{r}_j(s)}^2 \simeq \frac{\bar{\kappa}_{ij}(s) - \bar{r}_i(s)\bar{r}_j(s)}{N_s} \quad (i \neq j) \quad (\text{B.13})$$

$$\sigma_{\bar{r}_i(s)\bar{\kappa}_{ii}(s)}^2 \simeq \frac{2\bar{\kappa}_{ii}(s)}{tN_s} \quad (\text{B.14})$$

$$\sigma_{\bar{r}_j(s)\bar{\kappa}_{ij}(s)}^2 \simeq \frac{\bar{\kappa}_{ij}(s)}{tN_s} \quad (i \neq j) \quad (\text{B.15})$$

Note that the leading order fluctuations are those in the number of coincidences (i.e. $\sigma_{\bar{\kappa}_{ij}(s)}^2$). This provides a note of caution for measurers of correlation. For the estimate of the biases, which we are about to detail, in practice the variances may as well be computed numerically from the measurements as from these formulae; the formulae do however provide a check.

Carrying out the differentiation, we obtain for the bias of the three separate components of I_{tt} (denoted by $I_{tt1(\text{bias})}$, $I_{tt2(\text{bias})}$, $I_{tt3(\text{bias})}$):

$$\begin{aligned}
I_{tt1(\text{bias})} &= \frac{1}{\ln 2} \sum_s \widehat{\sum}_i \left[P(s) \ln(1/(1 + \nu_{ii})) \right. \\
&+ P(s)^2 \left(-1 + 2 \widehat{\sum}_j \frac{\bar{r}_j(s)}{\bar{r}_i} + 2 \frac{\bar{r}_i(s)}{\bar{r}_i} - \frac{\bar{r}_i^2(s)}{\langle \bar{r}_i^2(s') \rangle_{s'}} \right. \\
&\left. \left. - \widehat{\sum}_{j < \bar{r}_i(s') \bar{r}_j(s') >_{s'}} \frac{\bar{r}_j^2(s)}{\bar{r}_i^2} - \widehat{\sum}_j \frac{\langle \bar{r}_i(s') \bar{r}_j(s') \rangle_{s'}}{\bar{r}_i^2} \right) \right] \sigma_{\bar{r}_i(s)}^2 \quad (\text{B.16})
\end{aligned}$$

$$\begin{aligned}
I_{tt2(\text{bias})} &= \frac{-1}{\ln 2} \sum_s \widehat{\sum}_i \left[P(s) \left(\ln(1/(1 + \nu_{ii})) - 1 \right) \right. \\
&+ P(s)^2 \left(2 \widehat{\sum}_j \frac{\bar{r}_j(s)}{\bar{r}_i} + 2 \frac{\bar{r}_i(s)}{\bar{r}_i} - \frac{\bar{r}_i^2(s)}{\langle \bar{r}_i^2(s') \rangle_{s'}} \right. \\
&\left. \left. - \widehat{\sum}_{j < \bar{r}_i(s') \bar{r}_j(s') >_{s'}} \frac{\bar{r}_j^2(s)}{\bar{r}_i^2} - \widehat{\sum}_j \frac{\langle \bar{r}_i(s') \bar{r}_j(s') \rangle_{s'}}{\bar{r}_i^2} \right) \right] \sigma_{\bar{r}_i(s)}^2 \\
&+ \sum_s \widehat{\sum}_i \frac{P(s)}{\ln 2} \left[-\frac{P(s)}{\bar{r}_i^2} \widehat{\sum}_j \langle \bar{\kappa}_{ij}(s') \rangle_{s'} - \frac{\langle \bar{\kappa}_{ii}(s') \rangle_{s'}}{\langle \bar{r}_i^2(s') \rangle_{s'}} \right. \\
&+ P(s) \widehat{\sum}_j \frac{\langle \bar{\kappa}_{ij}(s') \rangle_{s'} \bar{r}_j^2(s)}{\langle \bar{r}_i(s') \bar{r}_j(s') \rangle_{s'}^2} + P(s) \frac{\langle \bar{\kappa}_{ii}(s') \rangle_{s'} \bar{r}_i^2(s)}{\langle \bar{r}_i^2(s') \rangle_{s'}^2} \left. \right] \sigma_{\bar{r}_i(s)}^2 \\
&+ \sum_s \widehat{\sum}_i \frac{2P(s)^2}{\ln 2} \left(\frac{1}{\bar{r}_i} - \frac{\bar{r}_i(s)}{\langle \bar{r}_i^2(s') \rangle_{s'}} \right) \sigma_{\bar{r}_i(s) \bar{\kappa}_{ii}(s)}^2 \\
&+ \sum_s \widehat{\sum}_{i \neq j} \frac{2P(s)^2}{\ln 2} \left(\frac{1}{\bar{r}_j} - \frac{\bar{r}_i(s)}{\langle \bar{r}_i(s') \bar{r}_j(s') \rangle_{s'}} \right) \sigma_{\bar{r}_j(s) \bar{\kappa}_{ij}(s)}^2 \quad (\text{B.17})
\end{aligned}$$

$$\begin{aligned}
I_{tt3(\text{bias})} &= \frac{1}{2 \ln 2} \sum_s \widehat{\sum}_i \left(\frac{P(s)}{\bar{\kappa}_{ii}(s)} - \frac{P(s)^2}{\langle \bar{\kappa}_{ii}(s') \rangle_{s'}} \right) \sigma_{\bar{\kappa}_{ii}(s)}^2 \\
&+ \frac{1}{\ln 2} \sum_s \widehat{\sum}_{i < j} \left(\frac{P(s)}{\bar{\kappa}_{ij}(s)} - \frac{P(s)^2}{\langle \bar{\kappa}_{ij}(s') \rangle_{s'}} \right) \sigma_{\bar{\kappa}_{ij}(s)}^2 \quad (\text{B.18})
\end{aligned}$$

The ‘hats’ in the summations of terms proportional to $\sigma_{\hat{r}_i(s)}^2$ have the same meaning as before; those in the summations of terms proportional to $\sigma_{\hat{r}_{ij}(s)}^2$ similarly restrict the sum to cases with non-zero underlying probabilities of observing coincidences from cells i, j in response to stimulus s (and similarly for the other terms).

Note that the leading contribution to the I_{tt} bias is from I_{tt3} , which is proportional to $1/t^2$, whereas the biases in I_{tt1} and I_{tt2} are only proportional to $1/t$.

We now have analytical expressions for the bias due to finite sampling in each of the components of I_{tt} , as well as I_t . The bias estimate obtained from each of these is subtracted from the ‘raw’ quantity. A more detailed study of the range of validity of the bias removal using simulated data can be found in (Schultz, 1998).

We conclude by noting that the procedure used to count ‘bins’ for the summation over ‘relevant bins’ in the above equations was a ‘naive’ counting procedure, in which we only add terms in which there is at least one spike (or coincidence if it is a sum over i and j) in any of the trials. For sufficiently short time windows, and a small number of trials per stimulus, the bias correction fails. This occurs more evidently in I_{tt3} because of the $1/t^2$ dependence. Other non-naive counting procedures can be used to obtain more accurate estimates of the bias. By using a Bayesian counting procedure (exactly the same one described in (Panzeri and Treves, 1996)), it was possible to reduce somewhat the time at which the bias correction broke down, and obtain a more accurate bias estimation at very short times, at the expense of losing the property that the resulting information estimate is an upper bound on the information. In this procedure, the problem of the summations reduces to estimating the number of relevant bins. This is done by choosing a “guess” value for the number of relevant bins, and a prior probability function which has one (constant) value for each of the occupied bins and another (constant) value for each of the empty bins.

The posterior probability distribution is then calculated and the posterior estimate of the number of relevant bins obtained. This procedure applies just as well to I_{tt3} as it did to the full information in the case described by Panzeri and Treves (1996). We do not describe this Bayesian counting procedure here, because it is exactly the one reported in Panzeri and Treves (1996); however, for values of firing rates in the range relevant for visual cortical cells, the use of a Bayesian counting procedure makes a difference only for time windows as short as 2-3 ms.

References

- Bevington, P. R. and Robinson, D. K. (1992). *Data reduction and error analysis for the physical sciences*, McGraw-Hill, New York.
- Panzeri, S. and Treves, A. (1996). Analytical estimates of limited sampling biases in different information measures, *Network* **7**: 87–107.
- Schultz, S. R. (1998). Information encoding in the mammalian cerebral cortex. D.Phil. Thesis, Oxford University.