Available online at www.isas.org.in/jisas

Journal of the Indian Society of Agricultural Statistics 68(2) 2014 117-129

Some Analyses of the Interaction among Local Field Potentials and Neuronal Discharges in a Mouse using Mutual Information

Apratim Guha

Indian Institute of Management, Ahmedabad, India

Received 23 July 2013; Revised 13 November 2013; Accepted 17 November 2013

SUMMARY

Constructing models for neuroscience data is a challenging task, more so when the data sets are of hybrid nature, and there exists very little work. As a first step, here we introduce a technique based on mutual information to look at bivariate hybrid time series data from the field of neuroscience. As an example, we use a data set on the local field potentials (which is a continuous time series) and nerve cell firings (which is a point process) of anaesthetized mice. We explore data driven confidence bounds for the mutual information statistics and discuss a test of independence between the two components of the hybrid process. A comparative study with the findings from some spectral domain methods are also discussed. It is found that the mutual information, as a time domain tool, complements the spectral domain methods.

Keywords: Hybrid process, Time series, Point process, Mutual information, Independence of time series components, Coherence, Phase.

1. INTRODUCTION

A possible way to differentiate among mice of different species, or same species but different genotypes is to compare the dynamics of the neurons in the auditory cortex, see Stiebler *et al.* (1997) for a discussion. In this paper, we aim to study tools that capture the inter-relationship between spike trains emanating from neurons in the temporal cortex of mice, which is a part of the auditory cortex, and local field potentials in the vicinity.

Data as above are clearly multi-dimensional and stochastic in nature, where the local field potentials are continuous in nature, and the firings can be viewed as point processes. Such multi-component stochastic processes, with some components time series and the rest point processes, are known as hybrid processes.

Although not many detailed studies of hybrid models exist, examples of hybrid situations in neuroscience are numerous. Bryant and Segundo (1976) characterizes the relation between transmembrane current (a time series) and the firing of voltage impulses (a point process) by a nerve cell of the sea hare Apysia Californica. Willie (1982a, b) discusses the relation between transmembrane current (a time series) and the firing of voltage impulses (a point process) by a nerve cell. Studies of biological systems often result in observation of hybrid processes. There may be continuously varying waveforms or sequences of discrete events (Halliday *et al.* 1995). In a more recent work, Brown *et al.* (1998) modelled spike trains jointly with the location of a rat, treating the location of the rat as a bivariate Gaussian random vector. See Valentine and Eggermont (2001) and Noreña and Eggermont (2002) for some other examples in the neuroscience field that deal simultaneously with spike trains and local field potentials. For various examples of hybrid situations, see Guha and Biswas (2008). Some

E-mail address : apratim@iimahd.ernet.in

examples and references could also be found in Amjad *et al*. (1997) and Halliday *et al*. (1999).

Correlation techniques are widely used in the time domain for study of linear dependence between signals, both in the time domain as well as the frequency domain. The examples are too many to list. However, when the dependence is of a non-linear nature, these linear methods are less effective. The need is of a suitable measure that captures the intrinsic non-linearity of the structure, towards which we employ an information theoretic tool called mutual information (MI), which, for two random variables X and Y is defined by

$$H(X,Y) = \int_x \int_y \log\left(\frac{p_{XY}(x, y)}{p_X(x) p_Y(y)} \right) p_{XY}(x, y) d\mu(x, y); \quad (1)$$

where μ is a suitable dominating measure, with respect to which the joint distribution of X and Y is differentiable. Notice that $H(X, Y)$ is 0 if X and Y are independent.

Note that the MI is a much more powerful technique that can be useful to investigate more general types of dependence structures, both linear and nonlinear. Our device is non-parametric and model free, and requires only some routine conditions to work.

In Section 2 some pertinent background material about some of the working principles of the auditory system of mammals, with particular attention to mice and include a description of the experimental setup. The dependence structure is assessed in the time domain using MI techniques.

2. DESCRIPTION OF THE STATISTICAL BACKGROUND

The local field potential (LFP), supposedly related to the so-called 'synaptic activity' of the neurons, is essentially recorded as a continuous time series. Following the discussions in Section 1, the LFP signals may be viewed as a stationary continuous time series, say $\{X(\cdot)\}$, and the number of spike discharges by a neuron from the beginning of the experiment is considered as a stationary orderly point process, say, $\{N(\cdot)\}$. Hence, the stationary discrete time series defined from $\{N(\cdot)\}$ by

$$Y(t) = N(t) - N(t - 1),$$

may be a 0-1 valued time series if the sampling interval is small.

Some stationarity assumptions may be used to justify the estimates. Local stationarity may be the all that is needed, but formally it is assumed, to make the analysis simpler, that both processes are strongly stationary. The assumption prima facie appears too strong but the time stretches here are brief and stationarity as an approximation is helpful. Further, there is no indication of non-stationarity in the data analysed in this paper.

To analyse the structure of the joint distribution of $\{X(t)\}$ and $\{Y(t)\}$, the MI may be used. For the remainder of this section, formulation of some pertinent MI estimates will be discussed.

The MI between $\{X(t)\}$ and $\{Y(t)\}$ at lag u, say $H(u)$, is given by

$$H(u) = \int_x \log\left(\frac{p(x,u)}{p_X(x) P(Y=1)} \right) p(x, u) dx + \int_x \log\left(\frac{p_X(x) - p(x,u)}{p_X(x) P(Y=0)} \right) (p_X(x) - p(x,u)) dx, \quad (2)$$

where u is a real number, for every fixed u, $p_{XN}(x, u)$ is positive and bounded away from 0 and ∞ uniformly on the range of $X(t)$, and

$$P[x < X(t) < x + dx] = p_X(x) dx;$$
$$P[dN(t) = 1] = p_{N1} dt \ \ \forall\, t. \quad (3)$$

To see how to get (2) from (1), note that

$$P[x < X(t) < x + dx, dN(t + u) = 1]$$
$$= P[x < X(t) < x + dx, N(t + dt + u) - N(t + u) = 1]$$
$$= p_{XN}(x, u)\, dx\, dt; \quad (4)$$

Notice that as $dt \to 0$, we have

$$P[x < X(t) < x + dx, \delta_{N(t+u)} = 1] = p_{XN}(x, u)\, dx$$

where $\delta_{N(t+u)}$ is the indicator of a spike at time $u + t$, which is effectively $Y(t + u)$ when the sampling intervals are small. Quite clearly,

$$P[x < X(t) < x + dx, \delta_{N(t+u)} = 0]$$
$$= P[x < X(t) < x + dx, Y(t + u) = 0]$$
$$= (p_X(x) - p_{XN}(x, u))dx.$$

One may alternatively view this idea simply as

$$f_{X(t)|Y(t+u)}(x|Y(t + u) = 1) \times P(Y(t + u) = 1) = p_{XN},$$

where $f_{X(t)|Y(t+u)}(\cdot)$ is the conditional density of $X(t)$ given $Y(t + u) = 1$.

Clearly two variables have zero MI if and only if they are independent, so that $\{X(t)\}$ and $\{Y(t)\}$ are

pairwise independent if and only if $H(u)$'s are zero for all real values of u. An estimate of $H(u)$ can therefore be used as a check for dependence at lag u by testing whether the estimate is significantly different from zero.

Next define

$$\hat{p}_X(x) = \frac{1}{Th_T}\sum_{t=1}^{T} K\left(\frac{X(t)-x}{h_T}\right)$$

$$\hat{p}_{XN}(x,u) = \frac{1}{Th_T}\sum_{t=1}^{T} K\left(\frac{X(t)-x}{h_T}\right) I_{N(t+u)-N(t+u-1)=1}, \quad (5)$$

where K is a kernel with finite support that is symmetric around zero, and integrates to 1. In our work here we employ the Epanechnikov kernel defined by

$$K(x) = \frac{3}{4}(1-x)^2 I_{|x|<1}.$$

Further, I_B is the indicator function of a set B defined as

$$I_B(x) = \begin{cases} 1 \text{ if } x \in B \\ 0 \text{ otherwise.} \end{cases}$$

Also, let

$$\hat{p}_{N1} = \frac{1}{T}\sum_{t=1}^{T} I_{N(t)-N(t-1)=1}; \quad (6)$$

$\hat{p}_{N1}$ is sometimes referred to as the empirical rate of the point process.

Now consider the test statistic

$$\hat{H}(u) = \int_X \log\left(\frac{\hat{p}_{XN}(x,u)}{\hat{p}_X(x)\hat{p}_1}\right)\hat{p}_{XN}(x,u)dx + \int_X \left(\frac{\hat{p}_X(x)-\hat{p}_{XN}(x,u)}{\hat{p}_X(x)(1-\hat{p}_1)}\right)(\hat{p}_X(x)-\hat{p}_{XN}(x,u))dx; \quad (7)$$

over all possible values of u, where the density estimates are as in (5-6). Then under the null hypothesis

$\mathbb{H}_0$: $\{X(t)\}$ and $\{Y(t)\}$ are pairwise independent, (8)

a large sample distribution for $\hat{H}(u)$ is developed in Guha (2005) under some regularity conditions. Since (8) can be restated as

$\mathbb{H}'_0$: H(u) = 0 for all $u \in \mathbb{R}$; (9)

one may utilize the large sample normality obtained by Guha (2005) to test $\mathbb{H}'_0$ against the alternative

$\mathbb{H}'_1$: $H(u_0) > 0$ for at least one $u_0 \in \mathbb{R}$; (10)

the same test also serves as a test for $\mathbb{H}_0$. However in practical examples we can only look at finitely many u, and hence one can consider $u \in \mathbb{Z} \bigcap [-n, n]$, for some large $n \in \mathbb{R}$, in (9) instead.

Note that if we reject the null hypothesis of independence for one lag, then we can reject the null hypothesis as a whole, so at the very least, we have a set of necessary but not sufficient tests of independence. However, given the high sampling rate of our data set, there should not be any practical problem with the lack of sufficiency of our tests, at least in the present context.

It is also pertinent to explore the dependence structure of the spike trains originating from the three neurons. Towards that goal, we discuss briefly some results to estimate the MI between two point processes.

Consider two orderly point processes $\{N_1(\cdot)\}$ and $\{N_2(\cdot)\}$ which are jointly stationary and the two corresponding 0-1 valued time series, $\{Y_1(\cdot)\}$ and $\{Y_2(\cdot)\}$, respectively. The MI between the two processes at lag u may be defined as

$$H_{Y_1Y_2}(u) = \sum_{i=0,1}\sum_{j=0,1} \log\left(\frac{p_{ij}(u)}{p_{1i}p_{2j}}\right) p_{ij}(u), \quad (11)$$

where $p_{ij}(u) = P(Y_1(t) = i,\ Y_2(t) = j)$ for $i, j = 0, 1$, $p_{ki} = P(Y_k(t) = i)$ for $k = 1, 2$; $i = 0, 1$.

Assume

$$0 < p_{11}(u) < \min(p_1^1, p_1^2); \quad (12)$$

$$\max(p_1^1, p_1^2) < 1. \quad (13)$$

When the two series are independent, $H_{Y_1Y_2}(u) = 0$ for all u.

The quantity $H_{Y_1Y_2}(u)$ can be estimated by a plug-in estimator $\hat{H}_{Y_1Y_2}(u)$ obtained by substituting the corresponding estimates of the probabilities in (11):

$$\hat{p}_{ij}(u) = \frac{1}{T-|u|}\sum_{t=\max(1,\,1-u)}^{\min(T-u,T)} I_{Y_1(t)=1,\,Y_2(t+u)=1} \quad (14)$$

$$\hat{p}_i^k = \frac{1}{T}\sum_{t=1}^{T} I_{Y_k(t)=j;}\ k = 1, 2; i = 0, 1. \quad (15)$$

The quantity in (14) can be simplified to $\frac{1}{T}\Sigma_{t=1}^{T} I_{Y_1(t)=1,\,Y_2(t+u)=1}$ when $|u| << T$.

Now, when the point processes are *M*-dependent the probability estimates as in (14- 15) will be almost surely (a.s.) consistent for the original probabilities; see Bosq (1996) for a similar result under weaker conditions.

The almost sure consistency of $\hat{H}_{Y_1Y_2}$ follows immediately when (12) is assumed. It is shown by Brillinger (2004) that $2T\hat{H}_{Y_1Y_2}(u)$ is approximately χ^2_1 under the assumption of independence, but the distribution changes to a Gaussian distribution when the two series are dependent.

3. A DESCRIPTION OF THE DATA

The resting membrane potential of the nerve cells serve as a signalling mechanism. When the resting potential is reduced beyond a certain threshold, usually about 10 mv, an all or none action potential is initiated. The action potential is a signal that lasts for a very short period, often 1 ms, could be as high as 110 mV and can be conducted at rates that vary between 1 and 100 meters per second (Kandel *et al*. (1991)). The information in the signals is coded in the frequency and number of spikes.

The local field potential (LFP) is supposedly related to the so-called 'synaptic activity' of the neurons; it is the compound electric field generated by many post-synaptic potentials in the vicinity (Asai *et al*. 2008). This means that even sub-threshold synaptic potentials may provide a contribution to LFP, in contrast to the spikes which are related exclusively to supra-threshold activities.

Hence one can think of the local field potential (LFP) as the total spiking activity of all units (neurons) within a certain region. However, in full generality this phrase actually refers to more: other sources that could generate electrical activities can not be disregarded, although the neurons are the main contributors.

For a more detailed description, see Kandel *et al.* (1991).

The data set analysed in this section was provided by Prof. A.E.P. Villa (University of Lausanne, Switzerland). It consists of three spike trains and one local field potential (LFP) series, collected by an electrode placed in the auditory cortex of an anaesthetized mouse. The three spike trains generated by three neurons are isolated from one electrode tip by a wave-shape sorting device.

Two experiments are considered here. Each of the tests lasts for approximately 100 seconds. The first experiment is associated with so-called spontaneous firing of the auditory neurons, whereas the last experiment involves the application of an auditory stimuli of band limited (or low-pass filtered) white noise bursts; (sound with a uniform frequency spectrum over a wide range of frequencies is referred to as white noise). The duration of the stimulus for this experiment is 200 milliseconds (ms), and the sampling rate of the LFP at a chosen location is 500 samples/sec (Hz). The firing times are recorded in the nearest ms.

Each of the spike train data files contains data for 99 "sweep"s, (a "sweep" is a period of time at the end of which the clock is reset to zero). The duration of each sweep was intended to be 1 second by the scientist, but actually they were always a few ms more than that. For example, for the spontaneous experiment, the first 99 spontaneous sweeps lasted 99.444 seconds instead of the intended 99 seconds; whereas for the stimulated case, the corresponding time was 99.348 seconds.

The LFP data were filtered after collection to eliminate the 50Hz component from the recordings. This was done because in Switzerland (where the experiments were conducted) the electric supply is provided with a frequency of 50Hz, and produces artificial peaks at 50Hz and its harmonics in the spectra. The filtering was performed to remove these misleading peaks from the data, but it was somewhat overdone and resulted in a notch at 50Hz in the resultant spectrum, see Fig. 1. A sharp peak at 100Hz and a small peak at 200Hz are still visible, which were not filtered out. The components for 0-7 Hz are also filtered out by the scientist.

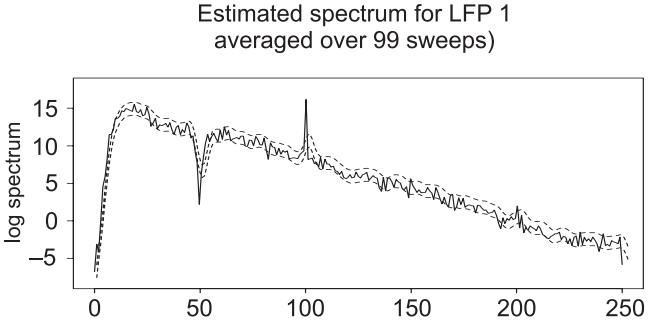

Fig. 1. The average estimated spectrum over the 99 sweeps of the LFP recordings for neuron 1. The broken line provides an approximate 95% confidence band around the spectrum.

3.1 Description of the Statistical Analysis

The primary objective of the analysis described in this section is to check for the possible existence of an interdependence (and possibly a causal relation) between the LFP recordings and the spikes, and to investigate possible changes in the dependence structure, if any, from spontaneous to stimulated experiments.

Data for each electrode, as described earlier in this section, consist of four series of observations: one series of local field potential (LFP) recordings and neuronal firings from three neurons.

Pairwise interrelationships among three spike trains and the local field potential recordings from a single neuron are investigated by estimating corresponding MIs. The spontaneous and the stimulated cases are compared. A spectral analysis is also carried out to substantiate the findings of the time domain analyses. We have already established in Section 2 that the LFP signals can be viewed as a stationary continuous time series, say $\{X(\cdot)\}$, and the number of spike discharges by a neuron over unit time periods can be considered as a stationary discrete 0-1 valued time series if the sampling interval is small. It was mentioned earlier in Section 3 that whereas the LFP recordings are recorded at 2 ms intervals, the spike firing times are recorded to the nearest millisecond. Hence, to make the data sets comparable, the spike firings are considered at 2 ms intervals. There exists a short time period after the firing of a neuron during which the neuron can not fire again, referred to as the refractory period of the neuron. It is observed that no neuron fires more than once within a time period of 4 ms, and hence it may be reasonable to assume that the refractory period of each neuron is greater than 2 ms; therefore the approximation does not contradict the consideration of the spike trains as 0-1 valued time series.

The rate of spikes is only about 1-2 per second and the data are collected for about 100 seconds. Because of this, the sum in $\hat{p}_{XN}(x,u)$ as in (5) involves too many terms which are zero, and hence may fail to provide reasonable estimates of the rate parameters. As a remedy, the estimators of the joint density function $p_{XN}(x, u)$ may be modified as follows:

$$\hat{p}_{XN}(x,u)=\frac{1}{2Th_Tb_T}\sum_{t=\max(1,\,b_T-u,\,b_T+u)}^{\min(T-u-b_T,\,T,\,T+u-b_T)}K\left(\frac{X(t)-x}{h_t}\right)\times\big((N(t+u+b_T)-N(t+u-b_T)\big); \qquad (16)$$

where $2b_T$, the bin-width for the point process counts, is a parameter asymptotically small compared to h_T. $\hat{H}(u)$ constructed using the definition of $\hat{p}_{XN}(x,u)$ either from (5) or (16) would have the asymptotic distribution as developed by Guha (2005).

One problem with the large sample estimates of standard errors of $\hat{H}(u)$ obtained from Guha (2005) is that the estimates could be poor even when based on moderately large data sets. As an alternative, standard error estimates may be calculated using blocked bootstrap methods.

The MI estimation for the spike trains suffer from the same problem of excess zeroes. We employ a strategy as above to estimate the MI for two spike trains as well, where we replace the estimates in (14) by

$$\hat{p}_{ij}(x,u)=\frac{1}{2Tb_T}\sum_{t=\max(1,b_T-u,b_T+u)}^{\min(T-u-b_T,T,T+u-b_T)}((N(t+b_T)-N(t-b_T))\times((N(t+u+b_T)-N(t+u-b_T)). \qquad (17)$$

One may note that in such a case, the asymptotic distribution suggested by Brillinger (2004) discussed earlier changes to $\chi^2_{(2b_T+1)^2}$, which can be approximated by a Gaussian distribution for a moderate or high value of b_T.

To calculate the confidence limits of $\hat{H}_{Y_1Y_2}(u)$, we may use bootstrapped estimates as before. We do not discuss the consistency issues of such bootstrapped estimators here, which may be subject to regularity conditions.

3.2 Methods and Results

In this section, blocked bootstrap based approximate 95% confidence intervals are presented, which are constructed for the hybrid MI, as given by (2), by taking the LFP and the spike trains into consideration as well as pairwise spike train MI, as given by (11), for the three neurons.

The bootstrap-based confidence intervals for the $H(u)$s, given by (2), are constructed under $\mathbb{H}_0$, achieved by sampling non-overlapping blocks of size 100 with replacement from both the series independently, and then computing the MI of the newly obtained series. This is repeated 100 times, and the approximate 95% confidence interval of the MI estimates at each lag, utilising the approximate normality obtained by Guha

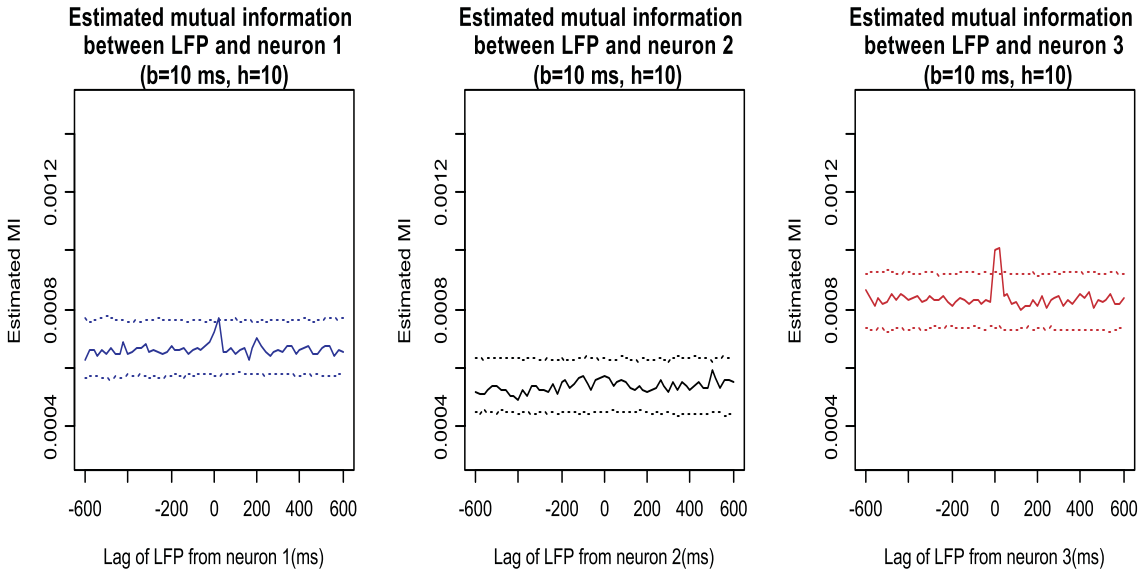

Fig. 2. Mutual information estimates between spontaneous firing recordings of the three neurons and the LFP series. The solid lines are MI estimates between local field potential recordings and neuronal firings; the dotted lines provide approximate 95% confidence intervals for the MI estimates, calculated by blocked bootstrapping.

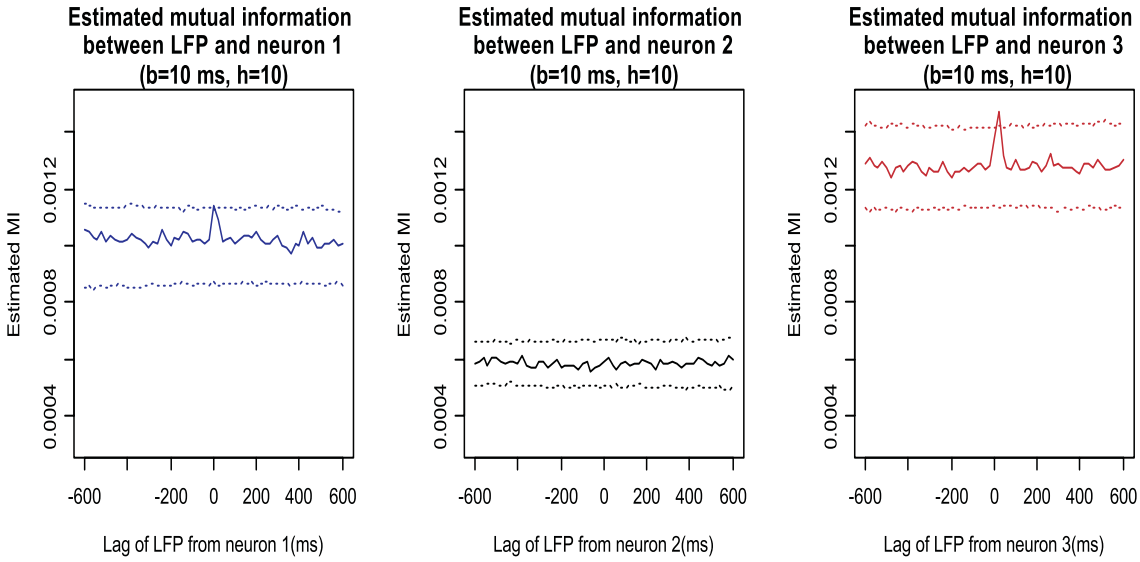

Fig. 3. Estimated MI between stimulated firing recordings of the three neurons and the LFP series. Approximate 95% confidence intervals calculated based on blocked bootstrap method are demarcated by the dotted lines.

(2005), is obtained by taking average $\hat{H}(u) \pm 2\widehat{s.e.}$ at each lag u. The results can be seen in Figs. 2 and 3.

Note that one may alternatively use the appropriate quantiles of the bootstrapped MI estimates as well. It is also possible to take the blocks without replacement; see the discussions towards the end of this section.

To calculate the confidence limits for the MIs between two spike trains which are formulated by (11) and estimated utilising (17), we employ 100 iterations of a blocked bootstrap with non-overlapping blocks of size 100 to construct the confidence limits for the MI estimates for the point process series under the assumption of independence of the two series. To estimate the MIs, we employ (16) and (17) with b_T = 50. As the distribution can now assumed to be nearly Gaussian by the discussion towards the end of Section 3.1, the confidence bounds are obtained, once again, using $\widehat{MI} \pm 2\widehat{s.e}$ at each lag u. The lower bounds are, however, not shown in the figures as they are almost 0 in every case.

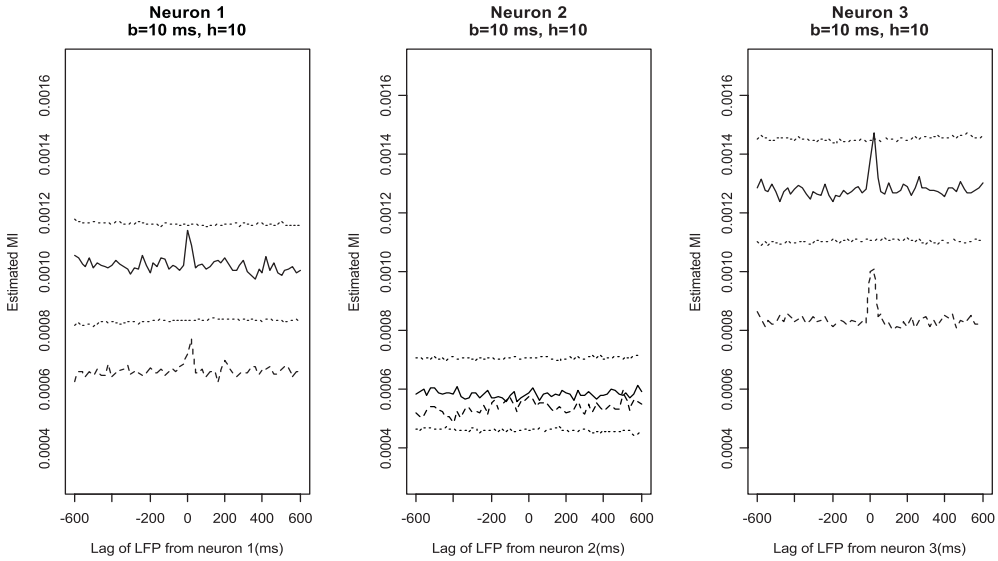

Fig. 4. Estimates of MI for the spontaneous and the stimulated cases are compared. The broken lines represent the estimates for the spontaneous case and the solid are the estimates for the stimulated case. Approximate 95% confidence intervals for the difference of the estimates, provided by the dotted lines, is shifted and plotted around the MI estimates for the stimulated case to provide a ready visual comparison.

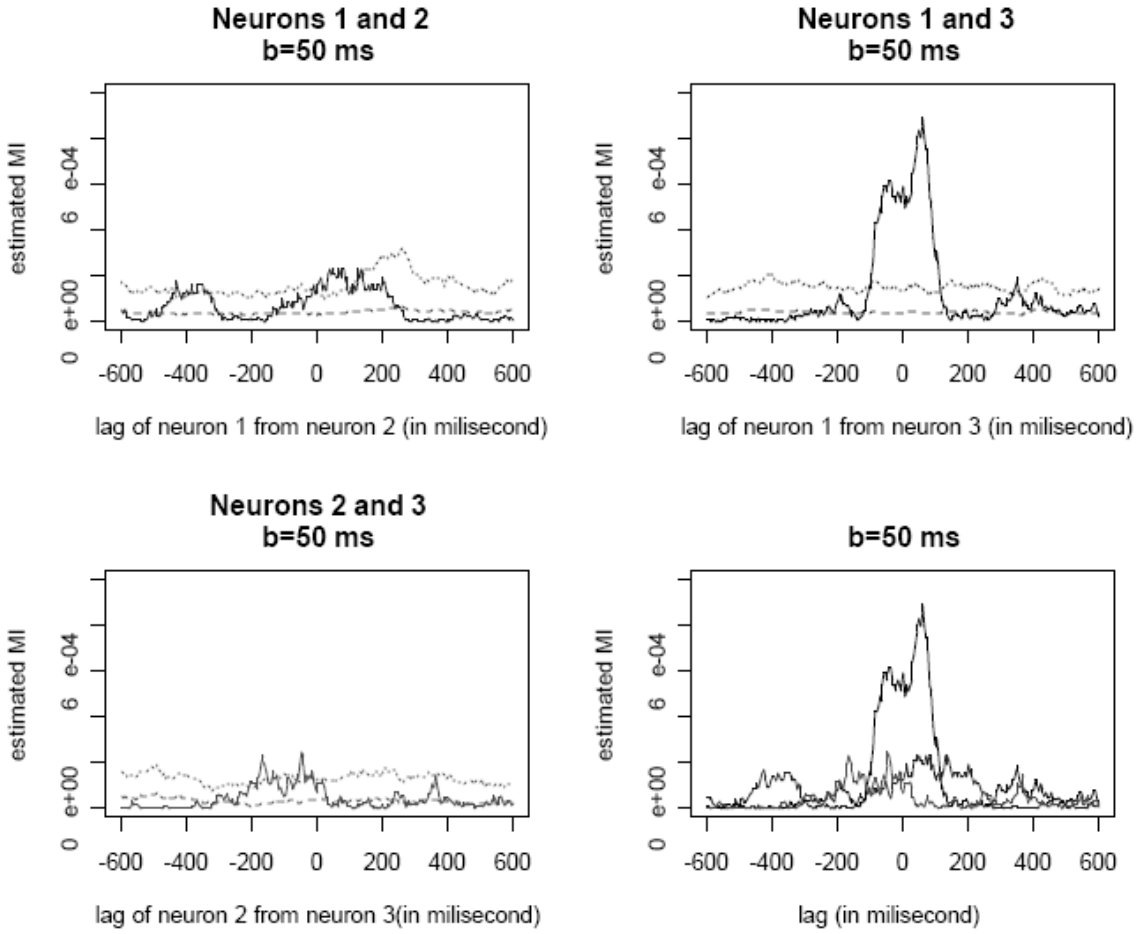

Fig. 5. Estimates of MI used to investigate the interrelationships among the three spike trains in the spontaneous case. The solid lines represent the mutual estimates. The dotted lines at the bottom and the broken line represent the estimated bias and approximate 95% confidence intervals based on blocked bootstrapping respectively.

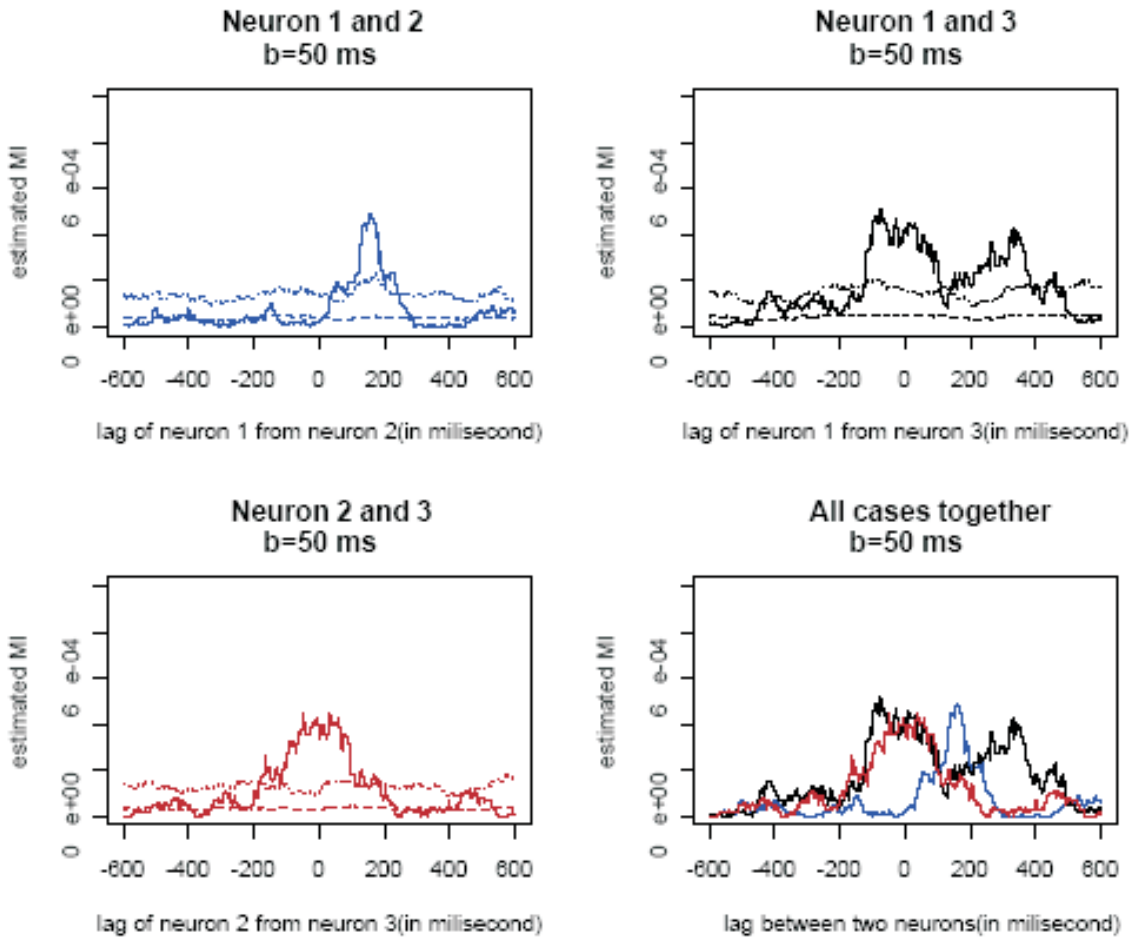

Fig. 6. Mutual information among the three spike trains in the stimulated case. Evidence of interrelationship is much stronger in all three cases.

It may be observed that the stimulated activities of the three neurons appear significantly related whereas only the first and third neuron seem to have any significant relationship among their spontaneous neuronal activities, see Figs. 5 and 6. The effect of the stimulus is apparently variable, for neurons 1 and 3 the stimulus seems to weaken the strength of the relationship; but it remains significant for a much longer period. On the other hand the relationship between neurons 2 and 3 is seemingly strengthened significantly by the introduction of a white noise stimulus.

The relationship between the LFP and the spike trains, the main issue of interest of the present analysis, seems very feeble in most parts from the hybrid MI estimates shown in Figs. 2 and 3. No evidence of any relationship between neuron 2 and LFP can be found. The MI between the spike train from neuron 1 and the LFP peaks for a time lag 0-40 ms of LFP from spike, so does the MI between the spikes from neuron 3 and the LFP; in the former case the MI are within the bootstrapped 95% confidence interval, whereas in the later, the observed peaks are outside the confidence intervals, but only slightly, so that it may be thought of being the effect of the 5% random variation not explained by the confidence interval.

Fig. 4 presents the spontaneous and stimulated cases together, and compares them visually. For each neuron, 95% confidence interval for the difference in hybrid MI over two situations is obtained through blocked bootstrap. If the confidence interval contains the zero line (*i.e.* the *x*-axis), that means at 5% level of significance we do not reject the null of equivalence. In the figure, we shift the confidence interval to be around the mean level of MI for the stimulated case, so that if the spontaneous MIs do not fall in that band, we reject the null, otherwise we do not. The plots clearly indicate that the stimulus increases the strength of relationship of the LFP series with the firings of neurons 1 and 3 significantly. The relation with neuron 2 is seemingly invariant and insignificant.

Before the peaks observed in the hybrid MI estimates in Figs. 2 and 3 are dismissed as random uctuations, it may be noted that in these figures, the confidence intervals seem to be too conservative when compared to the variation of the hybrid MI estimates away from the peaks. It is possible that this arises from multiple comparisons undertaken. Hence, we employ an alternative method to construct confidence regions to tackle this problem. In this method, non-overlapping blocks of the time series are sampled without replacement 19 times to obtain 95% confidence regions, for motivation see Brillinger and Guha (2007). The hybrid MI estimates thus calculated construct a 95% confidence region for MI between LFP and each of the three neurons, see Fig. 7 for the spontaneous situation and Fig. 8 for the stimulated situation. The graphs suggest that the first and third neuron firings are related with LFP. The hybrid MI estimates are significant for lags ~ 0-60 ms, so that it may be inferred that the firings of neuron 1 and 3 have causal effect on the LFP. Similar analysis is possible with the spike trains as well but is not shown here for brevity.

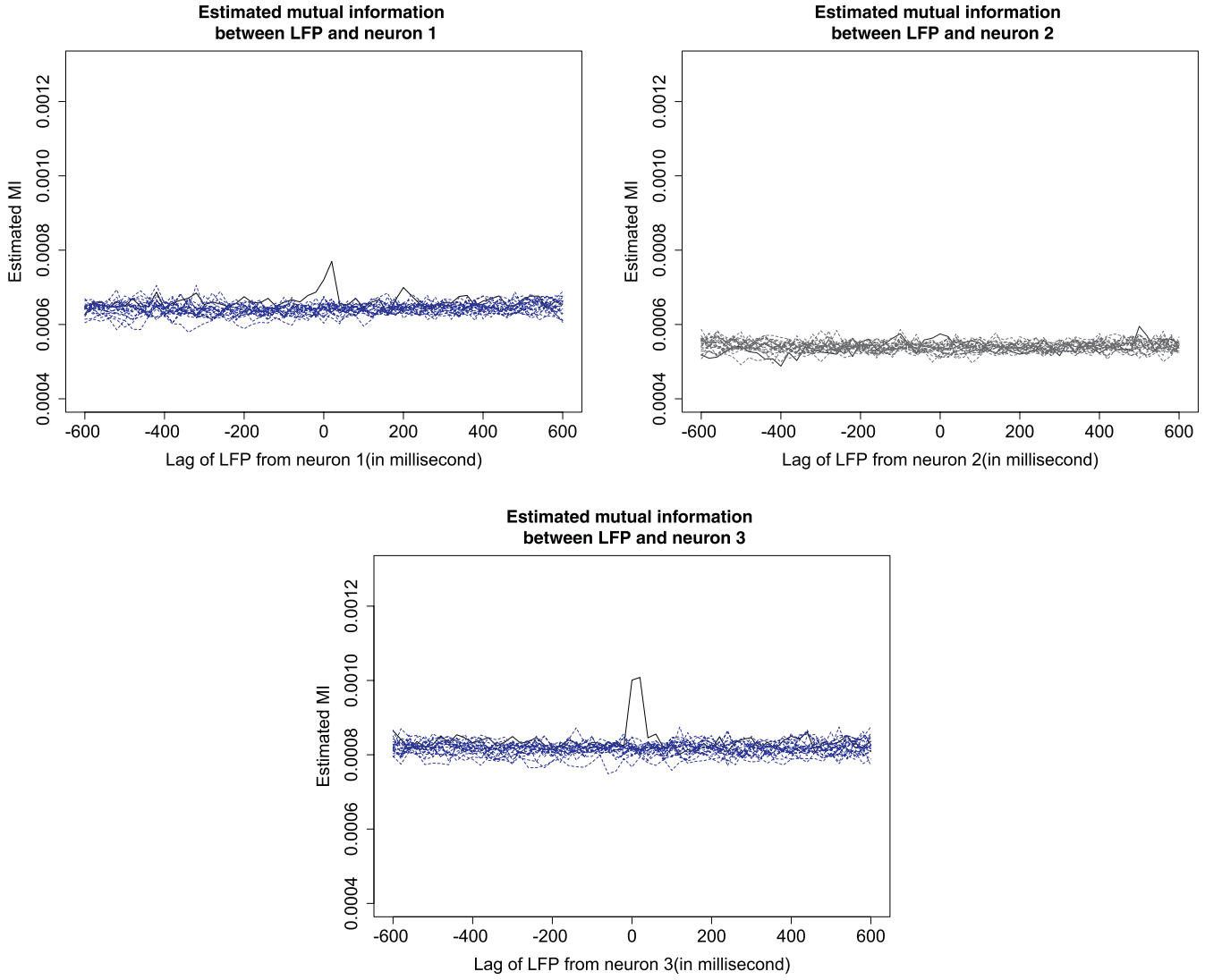

Fig. 7. Mutual information estimates between spontaneous firing recordings of the three neurons and the LFP series. The mesh of blue lines represent an approximate 95% confidence region obtained by calculating MI estimates for resampled series constructed by sampling blocks from the two original series independently without replacement.

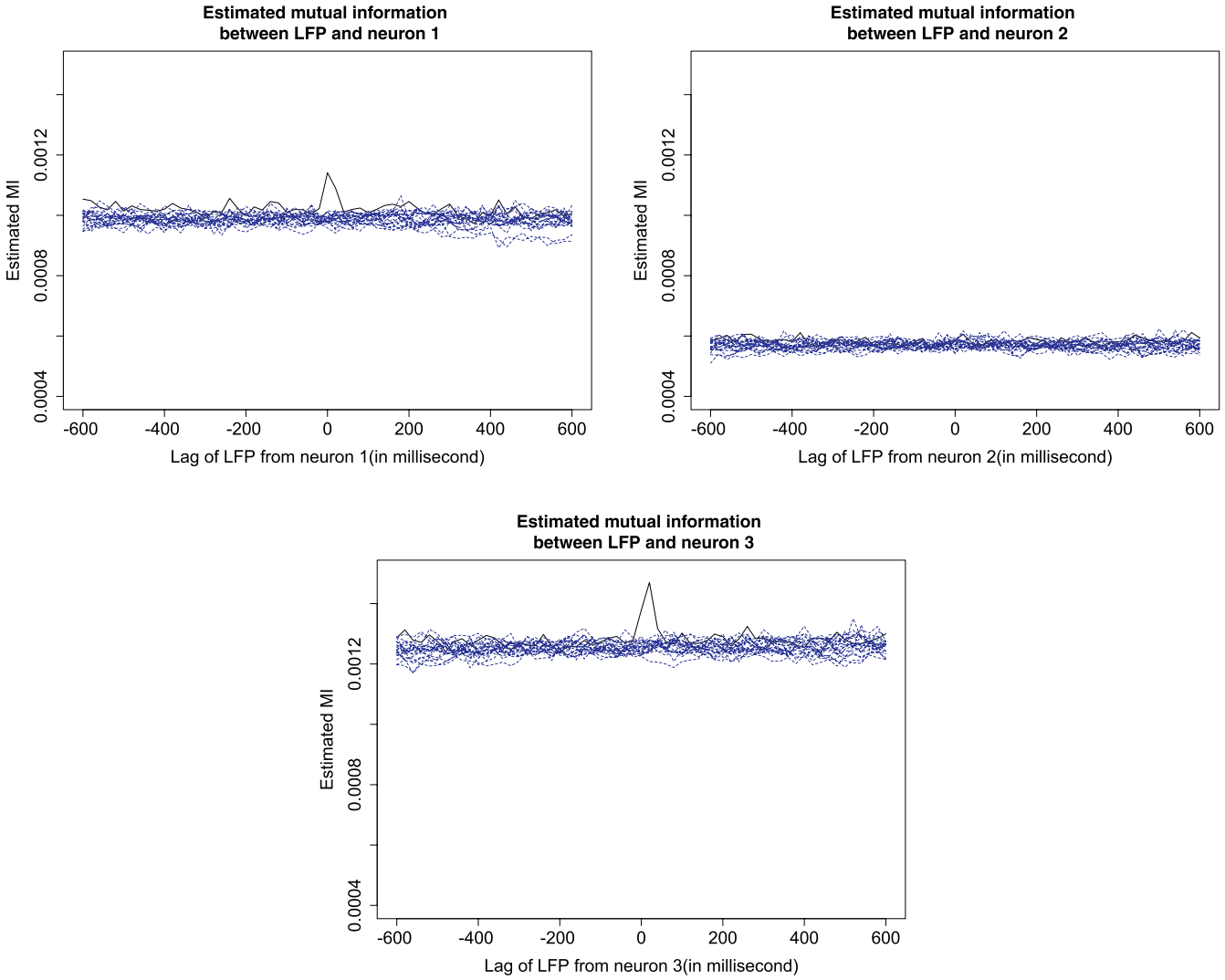

Fig. 8. Mutual information estimates between stimulated firing recordings of the three neurons and the LFP series. The mesh of blue lines represent an approximate 95% confidence region obtained by calculating MI estimates for resampled series constructed by sampling blocks from the two original series independently without replacement.

4. DISCUSSIONS

The analysis in this paper throws some light on the relationship between LFP and neuronal firings. The MI plots in Figs. 2 and 3 do not provide much evidence of causal relationship between LFP and neurons 2, but there is a hint of an association of LFP with neurons 1 and 3. The same observation is supported by Figs. 7 and 8.

To compare our findings with existing methods, we also obtained coherence and phase plots, definitions and motivation of which can be found in Brillinger (1975). The resulting estimates are shown in Figs. 9 and 10.

The coherence estimates between LFP and neuron 1 show some peaks around 75 Hz, which is sharper and higher in the stimulated case. Apart from a small peak near 25 Hz in the spontaneous case. The coherence estimate for neuron 2 does not reveal much information. Neuron 3 seems the most linked with LFP, with most of the frequencies appearing relevant.

Hence, the coherence estimates support the observations from the MI plots, but further suggests that the relation between LFP and neuron 3 to be stronger than that appears from MI.

On the basis of the MI and coherence estimates, neurons 1 and 3 have somewhat stronger association

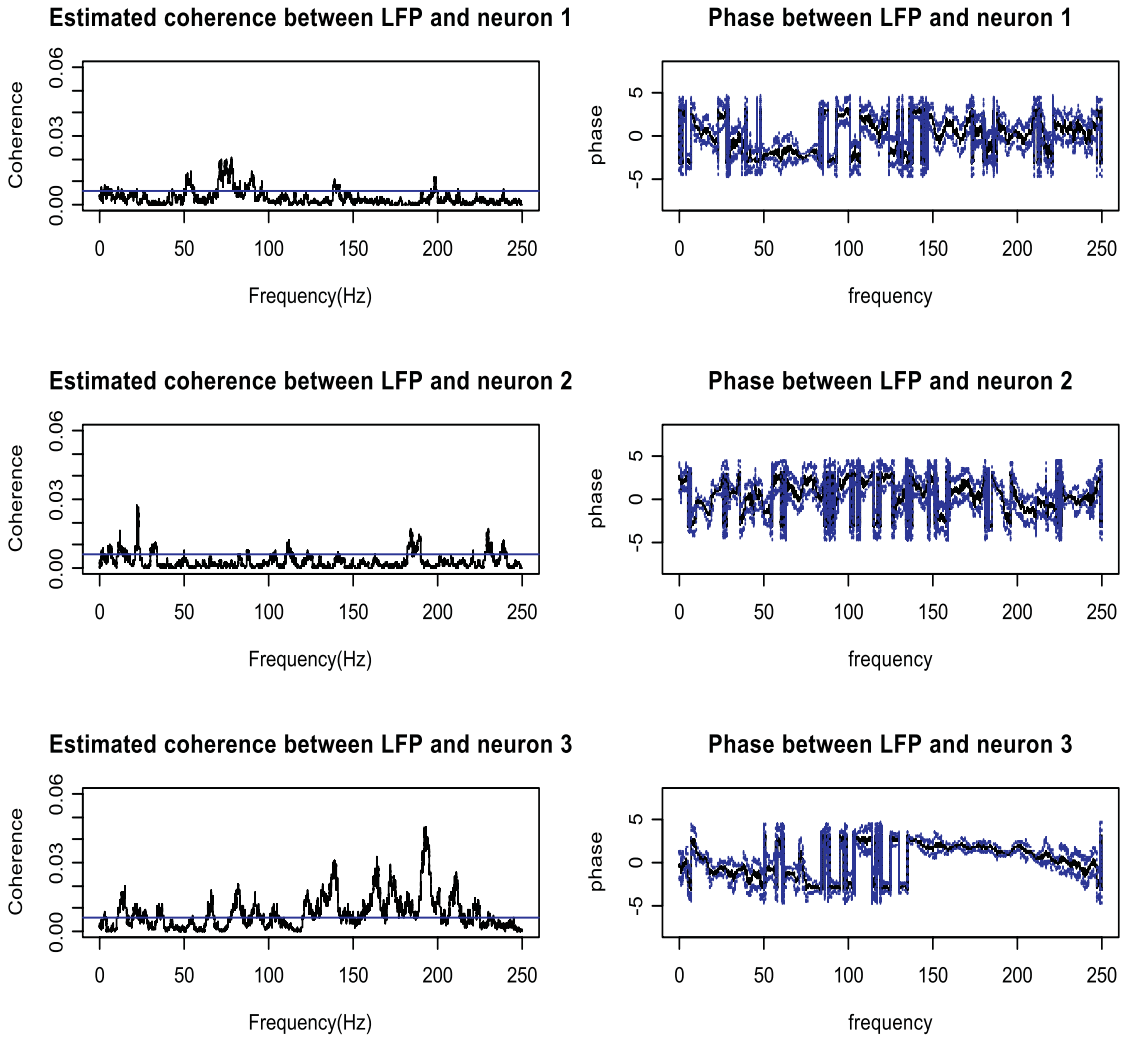

Fig. 9. Estimated coherence and phase between LFP and the three neuronal firings in the spontaneous case. The horizontal lines in the coherence estimate plots indicate approximate 95% upper confidence limit under the assumption of independence. For the phase plots the mesh indicate a 95% confidence interval.

with LFP when stimulus is applied. The MI seems to increase significantly, see Fig. 4. Sharper peaks are also observed in the coherence plot in Figs. 9 and 10.

Not much is revealed by the phase estimates between LFP and neurons 1 and 2 either. The phase estimates slope downwards for neuron 3, which may be indicative of existence of components of LFP that depend upon a delayed version of the neuronal spike trains for neurons 1 and 3. This finding may be exploited later in future studies, including an extensive study of possible effects of the spike trains on the LFP. Reverse effects, *i.e*. effects of the LFP on neuronal firings may also possibly exist, although not much evidence of such effects could be found in the analysis performed so far. That could be ascertained by further studies.

Beyond our findings for this particular example, however, we also note the utility of the MI techniques in general. Quite clearly they provide additional information on the time domain dependence structure that is not obtainable from the spectral analysis of the structure. Hence, it may be used as an additional tool, along with the conventional ones, to understand complex dynamic systems as in the example discussed here.

We note here that the coherence analysis is essentially based on application of correlation techniques in the frequency domain, and study only the linear relationships between the frequency components of the two series. It is possible to apply the MI techniques in the spectral domain as well to look at possible non-linear relationships among the various frequency components. For a motivating example see

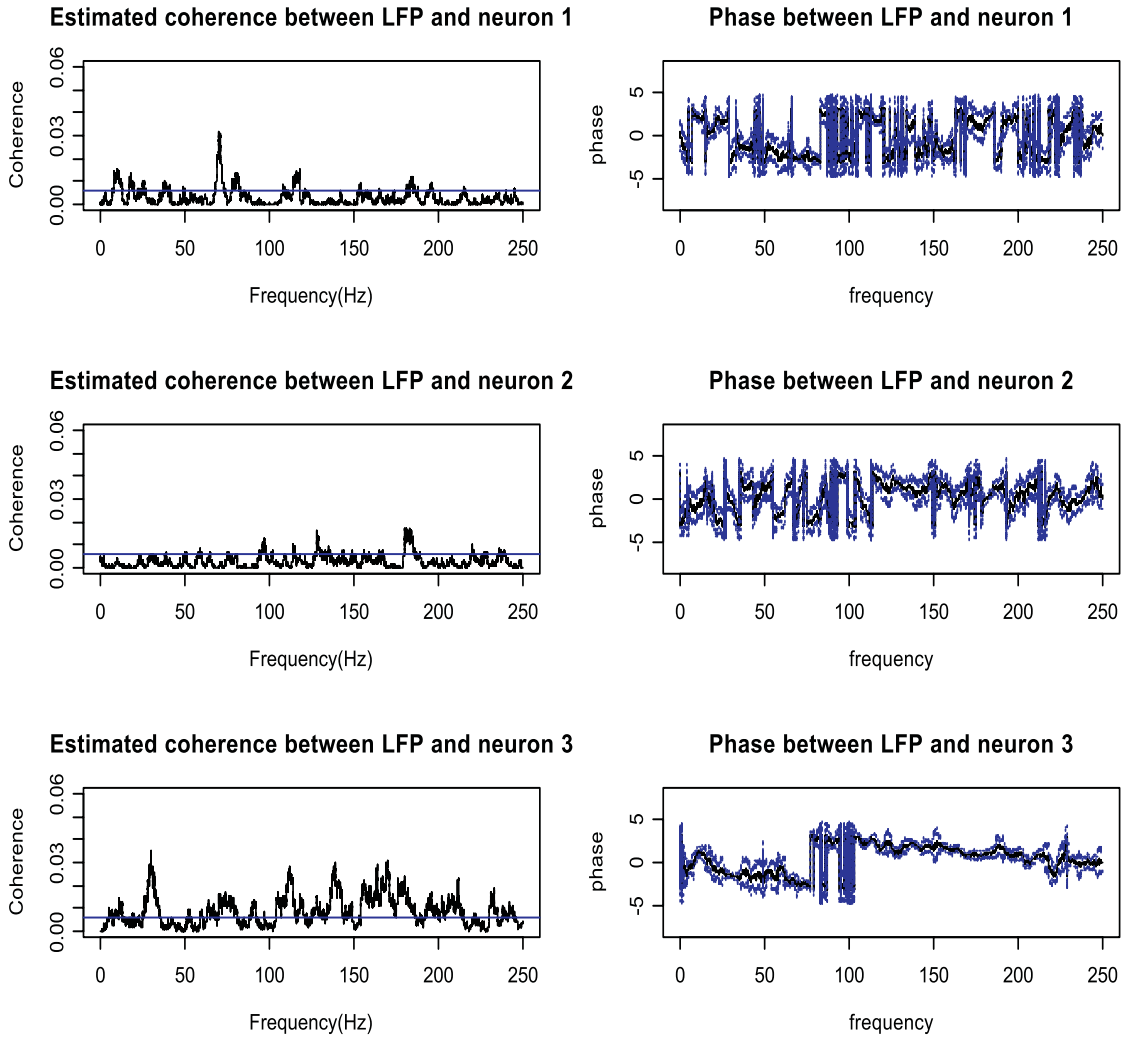

Fig. 10. Coherence and phase between LFP and the three neuronal firings when a white noise stimulus is applied. The horizontal lines in the coherence estimate plots indicate approximate 95% upper confidence limit under the assumption of independence. For the phase plots the mesh indicate a 95% confidence interval.

Brillinger and Guha (2007). However, such an analysis is computationally costly for data of the size of the present one, and hence was not taken up in the present study.

ACKNOWLEDGEMENTS

The author would like to thank two anonymous referees and the associate editor for many helpful suggestions which have lead to significant improvement of this paper.

Most of this work was carried out under the supervision of Prof. D.R. Brillinger (Department of Statistics, University of California, Berkeley) as a part of the author's PhD thesis.

The author also thanks Prof. A.E.P. Villa (Department of Information Systems, Faculty of Business and Economics, University of Lausanne, Switzerland) for kindly providing the data used in the analysis in this paper.

REFERENCES

Amjad, A.M., Halliday, D.M., Rosenberg, J.R. and Conway, B.A. (1997). An extended difference of coherence test for comparing and combining independent estimates-theory and application to the study of motor units and physiological tremor. *J. Neurosci. Methods*, **73**, 69-79.

Asai, Y., Guha, A. and Villa, A.E.P. (2008). Deterministic neural dynamics transmitted through neural networks. *Neural Networks,* **21,** 799-809.

Bosq, D. (1996). *Nonparametric Statistics for Stochastic Processes.* Springer-Verlag, New York.

Brillinger, D.R. (1975b). *Time Series Data Analysis and Theory*. Mcgraw Hill, San Francisco.

Brillinger, D.R. (2004). Some data analysis using mutual information. *Brazilian J. Probab. Statist.*, **18,** 163-182.

Brillinger, D.R. and Guha, A. (2007). Mutual information in the frequency domain. *J. Statist. Plann. Inf.,* **137**, 1076-1084.

Brown, E.N., Frank, L.M., Tang, D., Quirk, M.C. and Wilson, M.A. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal lace cells. *J. Neurosci.,* **18,** 7411-7425.

Bryant, H.L. and Segundo, J.P. (1976). Spike initiation by transmembrane current, a white noise analysis. *J. Physiol.,* **260,** 279-314.

Guha, A. (2005). Analysis of Dependence Structures of Hybrid Stochastic Processes using Mutual Information. Ph.D. Thesis, University of California, Berkeley.

Guha, A. and Biswas, A. (2008). An overview of modeling techniques for hybrid brain data. *Statistica Sinica*, **18,** 1311-1340.

Halliday, D.M., Rosenberg, J.R., Amjad, A.M., Breeze, P., Conway, B.A. and Farmer, S.F. (1995). A framework for the analysis of mixed time series/point process data-theory and application to the study of physiological tremor, single motor unit discharges and electromyograms. Progress in Bio-physics and Molecular Biology, **64,** 237-278.

Halliday, D.M., Conway, B.A., Farmer, S.F. and Rosenberg, J.R. (1999). Load-independent contributions from motor-unit synchronization to human physiological tremor, *J. Neurophysiol.,* **82,** 664-675.

Kandel, E.R., Schwartz, J.H. and Jessell, T.M., eds. (1991). *Principles of Neural Science.* 3rd edition, Appleton and Lange, Norwalk, Connecticut.

Noreña, A. and Eggermont, J.J. (2002). Comparison between local field potentials and unit cluster activity in primary auditory cortex and anterior auditory field in the cat. *Hearing Res.*, **166**, 202-213.

Stiebler, I., Neulist, R., Fichtel, I. and Ehret, G. (1997). The auditory cortex of the house mouse: left-right differences, tonotropic organization and quantitative analysis of frequency representation. *J. Comparative Physiol.,* **181,** 559-571.

Valentine, P.A. and Eggermont, J.J. (2001). Spontaneous burst-firing in three auditory cortical fields: its relation to local field potentials and its effect on inter-area cross-correlations. *Hearing Res.,* **154,** 146-157.

Willie, J. (1982a). Covariation of a time series and a point process. *J. Appl. Probab.,* **19,** 609-618.

Willie, J. (1982b). Measuring the association of a time series and a point process. *J. Appl. Probab.,* **19,** 597-608.